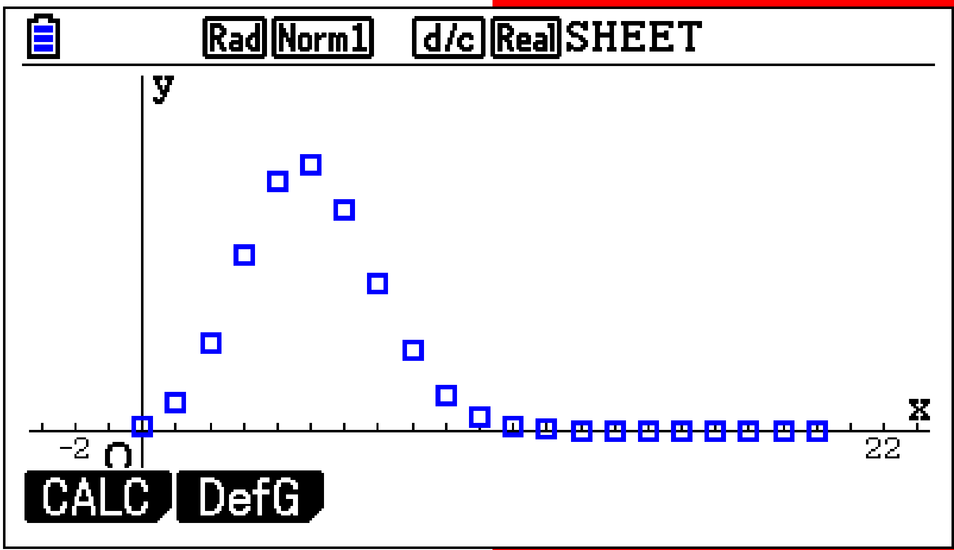
4. Unterrichtsvorhaben
in der Q1-Phase

Binomialverteilung



Jörn Meyer
j.meyer@fals-solingen.de
www.maspole.de

Inhaltsverzeichnis

1 Noch fit? – Stochastisches Grundwissen aus der E-Phase

Aufgabe 1: Wahrscheinlichkeitsverteilung und Erwartungswert

Definitionen: Den Ergebnissen eines Zufallsversuchs kann man **Wahrscheinlichkeiten** zuordnen. Die Wahrscheinlichkeiten sind gut gewählt, wenn sie die relativen Häufigkeiten bei großer Versuchszahl gut vorhersagen. Die Wahrscheinlichkeiten aller Ergebnisse addieren sich zu 100 %. Sie bilden eine **Wahrscheinlichkeitsverteilung**.

Beispiel: Im Zufallsversuch eines Wurfes von zwei Hexaeder-Würfeln erhält man folgende Wahrscheinlichkeitsverteilung für die Zufallsgröße X: *Anzahl 6er-Würfen*:

Ereignis	Zweimal 6	Einmal 6	Keine 6
Wahrscheinlichkeit	$\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}$	$\frac{5}{6}\cdot\frac{1}{6}+\frac{1}{6}\cdot\frac{5}{6}=\frac{10}{36}=\frac{5}{18}$	$1-\frac{11}{36}=\frac{25}{36}$

Definition: Wenn bei einer Datenerhebung die Ergebnisse $x_1, x_2, \ldots, x_n$ mit den Wahrscheinlichkeiten $p_1, p_2, p_3, \ldots, p_n$ auftreten, heißt der Wert $\mu = x_1 \cdot p_1 + x_2 \cdot p_2 + \cdots + x_n \cdot p_n$ der **Erwartungswert der Wahrscheinlichkeitsverteilung**. Er gibt an, welchen Mittelwert man bei ausreichend großer Versuchszahl auf lange Sicht erwarten kann. Er ist eine **Prognose für den Mittelwert**.

Beispiel: Wird beim obigen Zufallsversuch pro Wurf mit zwei Würfeln 1 € eingesetzt, stellt sich die Frage, wie die Auszahlung erfolgen muss, damit das Spiel fair ist. Betrachtet wird die folgende Auszahlungsvariante:

Ereignis	Zweimal 6	Einmal 6	Keine 6
Wahrscheinlichkeit	$\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}$	$\frac{5}{6}\cdot\frac{1}{6}+\frac{1}{6}\cdot\frac{5}{6}=\frac{10}{36}=\frac{5}{18}$	$1-\frac{11}{36}=\frac{25}{36}$
Auszahlung	8 €	1 €	0 €
Gewinn	7 €	0 €	- 1 €
Auszahlungserwartung	$\mu_A = 8\,€\cdot\frac{1}{36}+1\,€\cdot\frac{5}{18}+0\,€\cdot\frac{25}{36}=\frac{18}{36}€=0{,}50\,€$		
Gewinnerwartung	$\mu_G = 7\,€\cdot\frac{1}{36}+0\,€\cdot\frac{5}{18}+(-1\,€)\cdot\frac{25}{36}=-\frac{18}{36}€=-0{,}50\,€$		

Antwort: Das obige Spiel ist unfair, weil die Auszahlung auf Dauer nur 0,50 € pro Spiel ist, was zum einem dauerhaften Verlust von 50 Cent pro Spiel führen würde.

Begründe, für welchen Auszahlungsbetrag für einen 6er-Pasch, das Spiel fair wird.

Erst wenn z. B. die mittlere Auszahlung μ_A 1 € beträgt, kann das Spiel fair sein. Dafür verändert man zum Beispiel die Auszahlung für „Zweimal 6" auf a €. Es muss gelten:

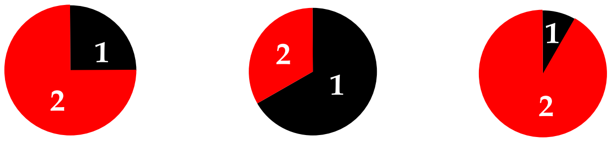
$$a\cdot\frac{1}{36}+1\cdot\frac{5}{18}+0\cdot\frac{25}{36}=1 \Leftrightarrow \frac{1}{36}a = 1-\frac{5}{18} \Leftrightarrow \frac{1}{36}a=\frac{13}{18} \Leftrightarrow a = 26$$

Werden für einen 6er-Pasch 26 € ausgezahlt, handelt es sich um ein faires Spiel.

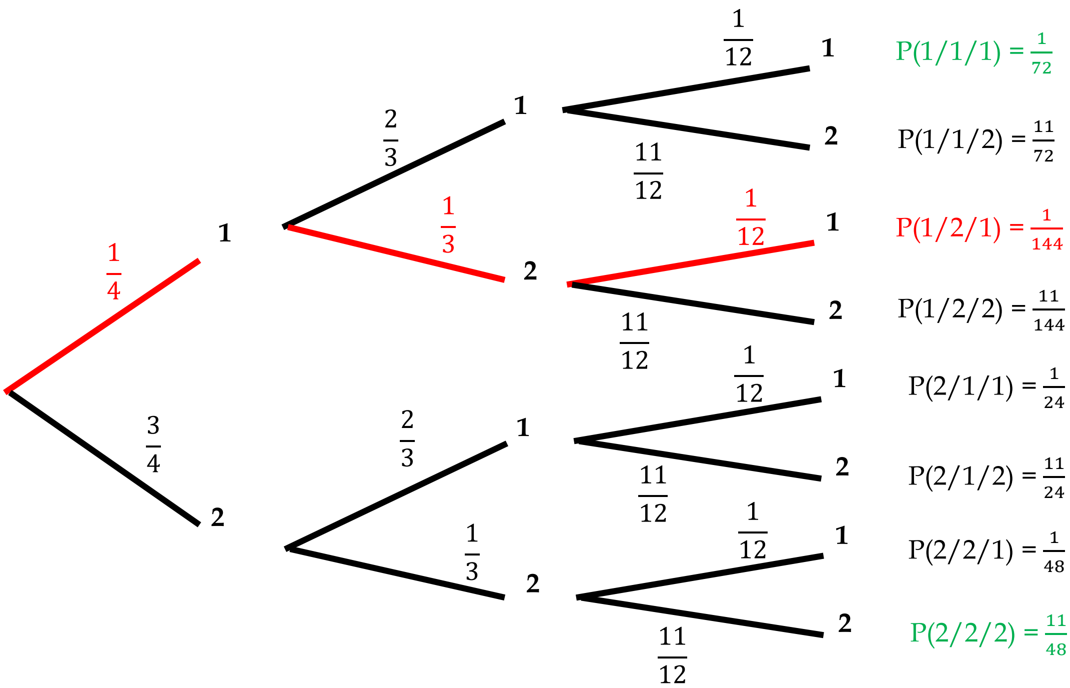
Aufgabe 2: Pfadregeln

Auf einem Jahrmarkt wird ein Glücksspiel angeboten. Für 1 € Einsatz werden nacheinander die folgenden 3 Glücksräder gedreht. Die Auszahlung erfolgt nach folgender Regelung:

Anzahl der 1er	0	1	2	3
Auszahlung in €	1	0	2	10



Man kann zu dem Zufallsversuch folgendes Baumdiagramm erstellen:



Will man die Wahrscheinlichkeit für das Ereignis (1/2/1) bestimmen, lässt sich diese mithilfe der **Pfadmultiplikationsregel** errechnen: $P(1/2/1) = \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{12} = \frac{1}{144}$.

Will man die Wahrscheinlichkeit für das Ereignis „Zahlendreierpasch“ bestimmen, muss man mithilfe der **Pfadadditionsregel** P(1/1/1) und P(2/2/2) addieren:

$P(\text{Zahlendreierpasch}) = P(1/1/1) + P(2/2/2) = \frac{1}{72} + \frac{11}{48} = \frac{35}{144}$

Für das obige Spiel ergibt sich folgende **Wahrscheinlichkeitsverteilung**:

Anzahl der Räder mit einer 1	0	1	2	3
Wahrscheinlichkeit	$\frac{11}{48}$	$\frac{5}{9}$	$\frac{29}{144}$	$\frac{1}{72}$

a) **Zeige**, dass das Spiel unfair ist. $\mu_A = 1 \cdot \frac{11}{48} + 0 \cdot \frac{5}{9} + 3 \cdot \frac{29}{144} + 11 \cdot \frac{1}{72} = \frac{71}{72} < 1$

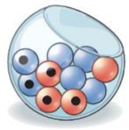
b) **Ermittle** eine Auszahlungsverteilung, die ein faires Spiel ermöglicht.

Das Spiel wird fair, wenn der Einsatz entweder $\frac{71}{72}$ € pro Spiel beträgt oder z. B. 12 € ausgezahlt werden, wenn 3 1er gedreht werden. Denn: $1 \cdot \frac{11}{48} + 0 \cdot \frac{5}{9} + 3 \cdot \frac{29}{144} + 12 \cdot \frac{1}{72} = 1$.

Aufgabe 3: Bedingte Wahrscheinlichkeit und Vierfeldertafel

Wenn man bei einer statistischen Erhebung zwei Merkmale wie z. B. Geschlecht und Körpergröße gleichzeitig untersucht, kann das Vorwissen über ein Merkmal die Wahrscheinlichkeit des anderen Merkmals beeinflussen. Man spricht von **bedingten Wahrscheinlichkeiten**.

Beispiel: Dies wollen wir am **Urnenmodell** verdeutlichen. Anhand der Abbildung rechts erkennt man, dass es vier unterschiedliche Kugelsorten gibt, dessen Anzahlen mithilfe einer **Vierfeldertafel** dargestellt werden können:

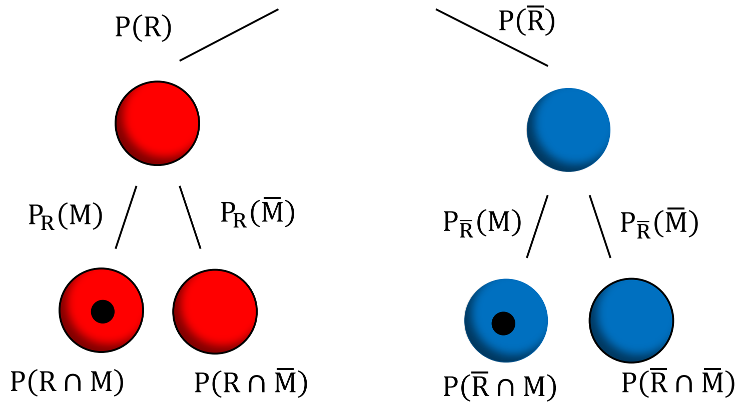


	Mit Punkt (M)	Ohne Punkt ($\overline{M}$)	Summe
Rot (R)	3	1	4
Nicht Rot ($\overline{R}$)	2	4	6
Summe	5	5	10

Dann beschreibt $\mathbf{P_R(M) = \frac{3}{4} = 75\,\%}$ die Wahrscheinlichkeit eine markierte Kugel (M) zu ziehen, wenn mann weiß, dass sie rot (R) ist. Von allen markierten Kugeln werden als Grundgesamtheit nur die roten Kugeln betrachtet. $\mathbf{P(R \cap M) = \frac{3}{10} = 30\,\%}$ beschreibt dagegen die Wahrscheinlichkeit, eine Kugel zu ziehen, die rot (R) und markiert ist (M). Von allen Kugeln werden solche betrachtet, die rot und markiert sind.

a) **Berechne** folgende Wahrscheinlichkeiten: $P_R(\overline{M}), P_{\overline{R}}(M), P_{\overline{R}}(\overline{M}), P_M(R), P_M(\overline{R}), P_{\overline{M}}(R), P_{\overline{M}}(\overline{R})$, die Wahrscheinlichkeiten $P(R \cap \overline{M}), P(\overline{R} \cap M)$ und $P(\overline{R} \cap \overline{M})$ sowie $P(R)$, $P(M)$ und $P(\overline{R})$ und $P(\overline{M})$. **Fasse** die Wahrscheinlichkeiten jeweils in **Worte**.

Die Vierfeldertafel kann zum Beispiel durch folgendes Baumdiagramme dargestellt werden:



Daraus lässt sich aufgrund der Pfadmultiplikationsregel folgender wichtiger Merksatz ableiten:

Merksatz: $\mathbf{P_A(B)}$ sei die bedingte Wahrscheinlichkeit für das Ereignis B, wenn man weiß, dass A eingetreten ist bzw. $\mathbf{P_B(A)}$ die entsprechende bedingte Wahrscheinlichkeit für das Ereignis A unter der Bedingung B. Dann gelten wegen der Pfadmultiplikationsregel folgende wichtige Formeln:

$$\mathbf{P_A(B) = \frac{P(A \cap B)}{P(A)}} \qquad \mathbf{P_B(B) = \frac{P(A \cap B)}{P(B)}}$$

b) **Erstelle** ein Baumdiagramm, das als erste Pfade danach unterscheidet, ob eine Kugel markiert ist oder nicht. **Beschrifte** alle Pfade mit den entsprechenden Wahrscheinlichkeiten.

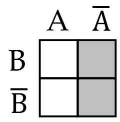
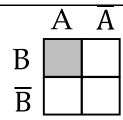
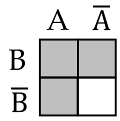
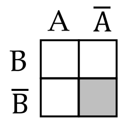
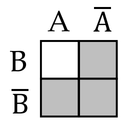
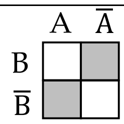
Wird etwa mit einem **Test** untersucht, ob jemand an einem Virus infiziert ist oder nicht, werden zwei Begriffe für bestimmte bedingte Wahrscheinlichkeiten verwendet:

Die **Sensitivität** entspricht der bedingten Wahrscheinlichkeit $P_{\text{infiziert}}(+)$ und bedeutet die Wahrscheinlichkeit, dass eine infizierte Person positiv getestet wird. Die **Spezifität** wird durch die bedingte Wahrscheinlichkeit $P_{\text{nicht infiziert}}(-)$ beschrieben und meint die Wahrscheinlichkeit, dass eine nicht infizierte Person einen negativen Test hat.

c) **Bestimme** die Sensitivität und Spezifität für folgenden Vierfeldertafel.

Wahrscheinlichkeiten	Infiziert (I)	Nicht infiziert ($\overline{I}$)	Gesamt
Positiver Test (+)	0,000999	0,001998	0,002997
Negativer Test (−)	0,000001	0,997002	0,997003
Gesamt	0,001 = 0,1 %	0,999	100 %

Aufgabe 4: Verknüpfungsmöglichkeiten von zwei Ereignissen

Sprechweise	**Term im mathematisches Modell**	**Veranschaulichung**
Gegenereignis zu A, Nicht das Ereignis A	$\overline{A}$	
Ereignis A und Ereignis B; Beide Ereignisse; Sowohl A als auch B	$A \cap B$	
Ereignis A oder Ereignis B; Mindestens eines der Ereignisse	$A \cup B = \overline{\overline{A} \cap \overline{B}}$	
Keines der Ereignisse; Weder A noch B	$\overline{A \cup B} = \overline{A} \cap \overline{B}$	
Höchstens eines der Ereignisse; Nicht beide Ereignisse	$\overline{A \cap B} = \overline{A} \cup \overline{B}$	
Genau eines der Ereignisse; Entweder A oder B	$(\overline{A} \cap B) \cup (A \cap \overline{B})$	

In der Saison 2014/2015 fielen in der Fußball-Bundesliga in 306 Spielen 843 Tore. 20 davon waren Eigentore. Aufgrund langfristiger Beobachtungen ist festgestellt worden, dass im Schnitt 55 % der Tore von **Stürmern** erzielt werden. 80 % aller Tore werden von Spielern erzielt, die in der **Startelf** stehen. Wenn ein **Einwechselspieler** ein Tor erzielt, ist dies zu 25 % **kein Stürmer**.

a) **Erstelle** zur dargestellten Situation eine geeignete Vierfelder-Tafel und **bestimme** dort die Anteile. **Markiere** die Anteile, die sich ohne Rechnung aus den obigen Angaben ergeben.

b) **Gib** Wahrscheinlichkeit und Ereignisterm für folgende Ereignisse an. Der Torschütze ist …

(1) Stürmer der Startelf.
(2) weder Stürmer noch Einwechselspieler.
(3) Stürmer oder Einwechselspieler.
(4) kein Stürmer.
(5) entweder Stürmer oder Einwechselspieler.
(6) nicht gleichzeitig Einwechselspieler und Nichtstürmer.

c) **Berechne** die Wahrscheinlichkeit, dass …

(1) unter den Stürmern ein Einwechselspieler getroffen hat.
(2) unter den Einwechselspielern ein Nichtstürmer getroffen hat.
(3) unter den Startspielern ein Stürmer getroffen hat.
(4) unter den Nichtstürmern ein Startspieler getroffen hat.

Aufgabe 5: Stochastische Unabhängigkeit

Definition: Zwei Ereignisse E und F heißen **stochastisch unabhängig**, wenn gilt: $\mathbf{P_E(F) = P(F)}$.

Satz: Zwei Ereignisse E und F sind genau dann stochastisch unabhängig, wenn die folgende Gleichung gilt: $\mathbf{P(E \cap F) = P(E) \cdot P(F)}$.

Man wirft zwei Würfel. **Untersuche** die Ereignisse A und B auf Unabhängigkeit.

a) A = „Die Augensumme ist 6.“ und B = „Die Differenz der Augenzahlen beträgt 0.“

b) A = „Der erste Würfel zeigt eine 3.“ und B = „Die Augensumme ist größer als 5.“

c) A = „Der erste Würfel zeigt eine Augenzahl unter 3.“ und B = „Der zweite Würfel zeigt eine Augenzahl über 3.“

2 Wahrscheinlichkeitsverteilung und Histogramme

Aufgabe 1: Differenz gewinnt

Spielanleitung:

- Jeder Spieler markiert 18 Striche in einer Farbe auf dem Spielfeld „Differenz", indem er sie auf die sechs Felder verteilt.
- Es wird im Uhrzeigersinn mit zwei Würfeln gewürfelt und die Differenz der Augenzahlen beider Würfel ermittelt.
- Vom Feld der gewürfelten Differenz darf jeder Spieler einen Strich seiner Farbe streichen.
- Gewonnen hat der Spieler, der zuerst alle seine Striche durchgestrichen hat.

a) Welche Strategie wählst Du? **Begründe** Deine Entscheidung? **Spiele** das Spiel und **überprüfe** Deine Ausgangsstrategie.

b) **Überlege**, welche Strategie am günstigsten ist und **fülle** die nachfolgende Tabelle Wahrscheinlichkeitsverteilung für die Zufallsgröße X: *„Differenz der Augensumme beim Wurf zweier Hexaeder"* **aus** und **zeichne** ein Histogramm. Hierfür führen wir einige wichtige **Definitionen** ein:

Zufallsgröße X: Eine **Zufallsgröße** oder **Zufallsvariable** X ist eine Funktion, die jedem Ergebnis eines Zufallsexperiments genau eine reelle Zahl x_i zuordnet. Im obigen Beispiel ordnet X jedem der 36 Möglichkeiten für einen zweifachen Hexaeder-Wurf genau eine der 6 möglichen Differenzen zu. Diese Differenzwerte werden mit $x_1 = 0$ $x_2 = 1$ …, $x_6 = 5$ bezeichnet.

Ereignis X = xi: Mit **X = xi** wird das **Ereignis** bezeichnet, dessen Ergebnisse alle dazu führen, dass die Zufallsgröße X den Wert x_i annimmt. Im obigen Beispiel sind dies X = 0 bis X = 5.

Wahrscheinlichkeitsverteilung und Histogramm: Ordnet man jedem Wert x_i, den die Zufallsgröße X annehmen kann, die Wahrscheinlichkeit $P(X = x_i)$ zu, erhält man eine Zuordnungstabelle, die man als **Wahrscheinlichkeitsverteilung** bezeichnet. Ihre grafische Darstellung heißt **Histogramm** oder Verteilungsdiagramm.

x_i	$x_1 = 0$	$x_2 = 1$	$x_3 = 2$	$x_4 = 3$	$x_5 = 4$	$x_6 = 5$
$X = x_i$	(1/1), (2/2) (3/3), (4/4) (5/5), (6/6)					
$P(X = x_i)$	$\frac{6}{36} = \frac{1}{6} = \frac{3}{18}$					
Histogramm $\frac{1}{18}$						

Zur Erinnerung: Der **Erwartungswert μ der Zufallsgröße X** mit der Wertemenge $\{x_1, x_2, \ldots, x_m\}$ entspricht in unserem Beispiel der mittleren Differenz, die man pro Hexaeder-2fach-Wurf erwarten kann. Es gilt: $\mu = E(X) = x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) + \cdots + x_m \cdot P(X = x_m)$. Er ist also eine Prognose für den Mittelwert.

c) **Berechne** den Erwartungswert E(X) für die obige Zufallsgröße X.

Spielfeld zu *„Differenz gewinnt“*

0	1
2	3
4	5

Arbeitsblatt 2: Augensummen

a) **Bestimme** die Wahrscheinlichkeitsverteilung der Zufallsgröße X: *„Augensumme beim 2-fachen Wurf eines Tetraeders"* und zeichne das zugehörige Histogramm.

x_i	$X = x_i$: zugehörige Ergebnisse	$P(X = x_i)$	Histogramm
2	(1/1)	$\frac{1}{16}$	
3			
4			
5			
6			
7			
8			

$\frac{1}{16}$

b) **Bestimme** die Wahrscheinlichkeitsverteilung der Zufallsgröße X: *„Augensumme beim Werfen eines Tetraeders und Hexaeders"* und zeichne das zugehörige Histogramm.

x_i	$X = x_i$: zugehörige Ergebnisse	$P(X = x_i)$	Histogramm
2			
3			
4			
5			
6			
7			
8			
9			
10			

$\frac{1}{24}$

c) Sowohl beim Werfen eines Oktaeder zusammen mit einem Tetraeder als auch beim Werfen von zwei Hexaedern sind die Augensummen 2, 3, ... , 12 möglich. **Bestimme** die dazugehörigen Wahrscheinlichkeitsverteilungen.

Woran könnte man bemerken, ob Oktaeder und Tetraeder bzw. zwei Hexaeder geworfen wurden?

x_i	**Oktaeder und Tetraeder**		**Zwei Hexaeder**	
	$X = x_i$	**$P(X = x_i)$**	**$X = x_i$**	**$P(X = x_i)$**
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				

d) Zur Erinnerung: Der **Erwartungswert μ** beschreibt, welcher Wert durchschnittlich bei einer großen Zahl von Würfen zu erwarten ist. Er ist eine Prognose für den Mittelwert.

Berechne für die vier Zufallsexperimente von Arbeitsblatt 2 jeweils den Erwartungswert.

Versuch	**Erwartungswert μ**
2-facher Wurf eines Tetraeders	
Wurf eines Tetraeders und Hexaeders	
Wurf eines Oktaeders und Tetraeders	
Wurf zweier Hexaeder	

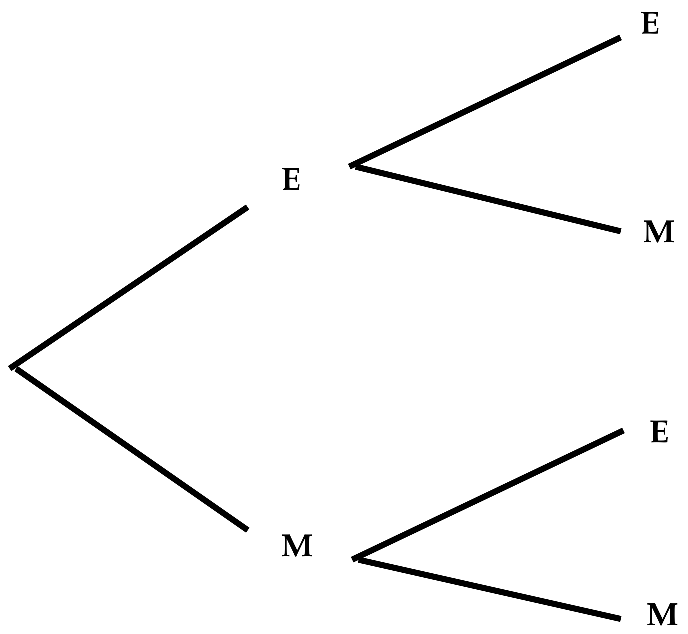
3 Binomialverteilung

Aufgabe 1 (Überraschungs-Schokokugeln)

Ein Hersteller von Überraschungs-Schokokugeln wirbt damit, dass in jeder siebten Kugel eine Figur enthalten ist, die unter Sammlern als besonders wertvoll gilt. Daher ist die Freude groß, wenn man in seiner Kugel eine solche Figur entdeckt. Ein solcher Kugelinhalt werde also als Erfolg (Treffer) gewertet (E), alle anderen Kugelfüllungen als Misserfolg (M).

Man betrachte zunächst den Kauf von zwei Überraschungskugeln und untersuche die Wahrscheinlichkeitsverteilung der Zufallsgröße X: *„Anzahl der Erfolge“* (d. h. Anzahl der Überraschungskugeln mit besonderem Inhalt).

a) **Trage** in folgendem Baumdiagramm die Pfadwahrscheinlichkeiten **ein**.



b) **Ermittle** die möglichen Ergebnisse und Wahrscheinlichkeiten für 0, 1, 2 Erfolge beim Kauf von zwei Schokokugeln und trage sie in die Tabelle ein.

Erfolgsanzahl k	**X = k: mögliche Ergebnisse mit k Treffern**	**Wahrscheinlichkeit P(X = k)**
0		
1		
2		

c) **Ergänze** das obige Baumdiagramm für den Kauf von frei Schokokugeln, **ermittle** die möglichen Ergebnisse und Wahrscheinlichkeiten für 0, 1, 2, 3 Erfolge und **trage** sie in die Tabelle **ein**.

K	X = k	P(X = k)
0		
1		
2		
3		

d) **Fülle** die Wahrscheinlichkeitsverteilung für vier Schokokugeln aus.

K	X = k	P(X = k)
0		
1		
2		
3		
4		

e) **Untersuche**, wie sich die Wahrscheinlichkeit P(X = k) für das Ereignis „k Erfolge" bei einer beliebigen Anzahl n gekauften Überraschungsschokokugeln berechnet.

f) Berechne den **Erwartungswert μ** für die Zufallsgröße X: „Anzahl k der Erfolge beim Kauf von n Überraschungsschokokugeln" für n = 1, 2, 3, 4 Schokokugeln.

N	Erwartungswert μ
1	$0 \cdot \frac{6}{7} + 1 \cdot \frac{1}{7} = \frac{1}{7}$
2	
3	
4	
n	

Aufgabe 2: Iterative Berechnung von Wahrscheinlichkeiten bei Bernoulli-Ketten

> **Definition:** Bei jedem Zug einer Schokokugel sind genau zwei Ausgänge möglich, nämlich Erfolg (E) und Misserfolg (M). Solche Zufallsversuche (Experimente) nennt man **Bernoulli-Experimente**.

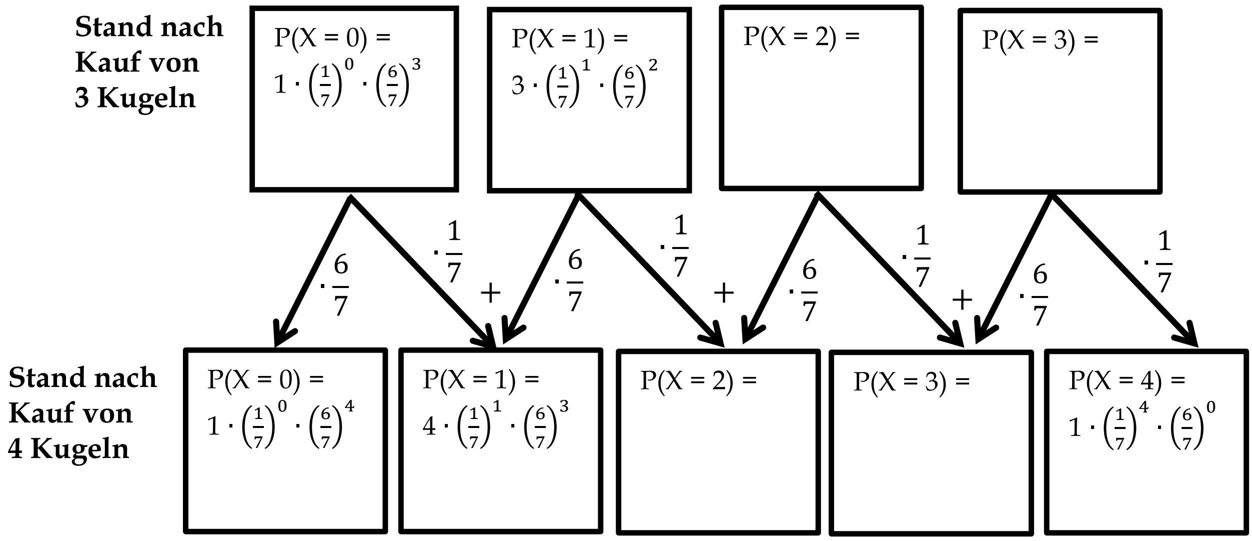
a) Nenne weitere Bernoulli-Versuche.

> **Definition:** Ändert sich die Erfolgswahrscheinlichkeit bei jedem der n Züge eines Bernoulli-Versuches nicht, spricht man von einer **Bernoulli-Kette der Länge n** (n-stufige Bernoulli-Kette).

b) **Gib** Beispiele für Bernoulli-Versuche an, bei denen sich die Erfolgswahrscheinlichkeit im Laufe eines n-stufigen Experimentes **ändert**.

c) Nach dem Kauf von Überraschungsschokokugeln kann man die Wahrscheinlichkeiten der Ereignisse auch anders aufschreiben. Für den Kauf von drei Kugeln liegt also eine **3-stufige Bernoulli-Kette mit der Erfolgswahrscheinlichkeit p = $\frac{1}{7}$** vor.

Notiere in den Kästchen in der Zeile „Stand nach 3 Käufen" Terme für die fehlenden Wahrscheinlichkeiten. **Erläutere** die Bedeutung der beschrifteten Pfeile und **gib** Terme für die fehlenden Wahrscheinlichkeiten in der Zeile „Stand nach 4 Kugeln" **an**.



d) **Erläutere**, wie man von den Wahrscheinlichkeiten der ersten zu denen der zweiten Zeile gelangt.

e) **Erweitere** das Diagramm für n = 5, indem Du Terme für Wahrscheinlichkeiten in die nachfolgenden Kästchen einträgst.

P(X = 0) = $1 \cdot \left(\frac{1}{7}\right)^0 \cdot \left(\frac{6}{7}\right)^5$	P(X = 1) =	P(X = 2) =	P(X = 3) =	P(X = 4) =	P(X = 5) =

f) **Erläutere**, wie man überprüfen kann, dass nichts übersehen wurde.

Berechnung der Pfadanzahl mit k Erfolgen führt zum Binomialkoeffizienten $\binom{n}{k}$

Will man z. B. wissen, wie viele Pfade bei 5 Schokokugeln zu 3 Erfolgen führen, überlegt man sich ein passendes Bild. Man könnte sich fragen, wie viele Möglichkeiten 3 Personen haben sich auf 5 Stühle zu setzen. Dabei darf die Reihenfolge (Individualität) der sitzenden Personen allerdings keine Rolle spielen. Denn bei einem Ergebnis (Erfolg, Erfolg, Erfolg, Misserfolg, Misserfolg) spielt die Reihenfolge (Individualität einer Schokokugel mit Erfolg) keine Rolle.

3 Personen haben offenbar $5 \cdot 4 \cdot 3 = 60$ Möglichkeiten sich auf 5 Stühle zu setzen. Denn: Die erste Person hat 5 Möglichkeiten sich zu setzen, die zweite zu jeder dieser Möglichkeiten 4 Möglichkeiten, so dass insgesamt $5 \cdot 4 = 20$ Möglichkeiten entstehen. Zu jeder dieser $5 \cdot 4 = 20$ Möglichkeiten hat eine dritte Person 3 Möglichkeiten, also insgesamt $20 \cdot 3 = 5 \cdot 4 \cdot 3 = 60$ Möglichkeiten.

Da die Reihenfolge der Stuhlbelegungen keine Rolle spielen soll, haben wir zu viele Möglichkeiten. Zu jeder der $5 \cdot 4 \cdot 3 = 60$ Möglichkeiten (**Permutationen**) führen jeweils $3 \cdot 2 \cdot 1 = 6$ Möglichkeiten zur gleichen **Kombination**. Im Bild gesprochen: Bei jeder Sitzkombination können sich die drei Personen mit $3 \cdot 2 \cdot 1 = 6$ Möglichkeiten auf die besetzten drei Stühle verteilen, ohne dass eine „neue" Kombination entsteht.

Insgesamt haben 3 Personen **ohne Beachtung der Reihenfolge** (Individualität) $\frac{5 \cdot 4 \cdot 3}{3 \cdot 2 \cdot 1} = \frac{60}{6} = 10$ Möglichkeiten, sich auf 5 Stühle zu setzen. **Beachtet** man die **Reihenfolge** (Individualität), erhält man $5 \cdot 4 \cdot 3 = 60$ Möglichkeiten.

Man schreibt im ersten Fall $\binom{5}{3} = \frac{5 \cdot 4 \cdot 3}{3 \cdot 2 \cdot 1} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = \frac{5!}{3! \cdot (5-3)!}$ und nennt diesen Ausdruck **Binomialkoeffizient 5 über 3**. Im zweiten Fall gilt $5 \cdot 4 \cdot 3 = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1} = \frac{5!}{(5-3)!}$.

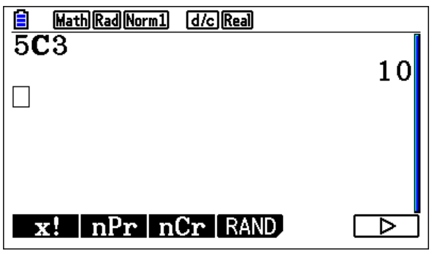
Verallgemeinert man die obige Überlegung auf n Stühle und k Personen, die sich ohne Beachtung der Individualität darauf verteilen, erhält man für den **Binomialkoeffizient n über k**:

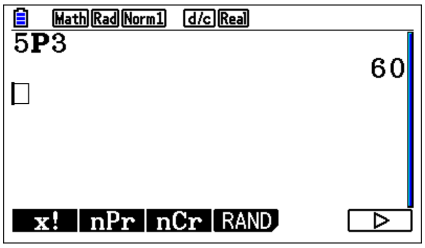
$$\binom{n}{k} = \frac{n \cdot (n-1) \ldots [n-(k-1)]}{k \cdot (k-1) \cdots 2 \cdot 1} = \frac{n\,(n-1) \ldots [n-(k-1)] \cdot (n-k) \cdots 2 \cdot 1}{k \cdot (k-1) \cdots 2 \cdot 1 \cdot (n-k) \cdots 2 \cdot 1} = \frac{\mathbf{n!}}{\mathbf{k! \cdot (n-k)!}}$$

Beachtet man die Reihenfolge der k Personen, sich auf n Stühle zu setzen, erhält man folgende Anzahl an Möglichkeiten:

$$n \cdot (n-1) \ldots [n-(k-1)] = \frac{n \cdot (n-1) \ldots [n-(k-1)] \cdot (n-k) \cdots 2 \cdot 1}{(n-k) \cdots 2 \cdot 1} = \frac{\mathbf{n!}}{\mathbf{(n-k)!}}$$

Mit der GTR kann der Binomialkoeffizient berechnet werden im $\boxed{\text{MENU 1}}$ über $\boxed{\text{OPTN}}$, $\boxed{\text{F6}}$, $\boxed{\text{F3}}$ und den Menüpunkt $\boxed{\text{nCr}}$ (*combination* = Kombination) sowie die Anzahl der Kombinationen unter Beachtung der Reihenfolge im Menüpunkt $\boxed{\text{nPr}}$ (*permutation* = Permutation):





g) **Berechne** folgende Ausdrücke händisch und mit dem GTR: $\binom{10}{5}$; $\binom{6}{4}$; $\binom{10}{10}$; $\binom{10}{1}$; $\binom{100}{99}$; $\frac{100!}{99!}$

h) **Erläutere** am Beispiel von k = 3 Frauen und n – k = 5 – 3 = 2 Männern, die auf n = 5 Stühle verteilt werden sollen, den Binomialkoeffizienten $\binom{5}{3}$.

i) Allgemein hängt die Art der iterativen Berechnung mit dem **Pascalschen Dreieck** zusammen. **Erläutere** das Schema und **beweise** die Umformungen im unteren Kasten.

k - 1 Erfolge bei n - 1 Versuchen:	k Erfolge bei n - 1 Versuchen:
$P_{n-1}(X = k - 1)$	$P_{n-1}(X = k - 1)$
$= \binom{n-1}{k-1} \cdot p^{k-1} \cdot (1-p)^{n-k}$	$= \binom{n-1}{k} \cdot p^{k} \cdot (1-p)^{n-1-k}$

ein Erfolg im n-ten Versuch, d. h. $\cdot\, \mathbf{p}$ **+** **ein Misserfolg im n-ten Versuch, d. h.** $\cdot\, \mathbf{1-p}$

k Erfolge bei n Versuchen:

$$\mathbf{P_n(X = k)}$$

$$= \left[\binom{n-1}{k-1} + \binom{n-1}{k}\right] \cdot p^k \cdot (1-p)^{n-k} = \boldsymbol{\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}}$$

$$\begin{array}{c}
\binom{0}{0} \\
\binom{1}{0} \quad \binom{1}{1} \\
\binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2} \\
\binom{3}{0} \quad \binom{3}{1} \quad \binom{3}{2} \quad \binom{3}{3} \\
\binom{4}{0} \quad \binom{4}{1} \quad \binom{4}{2} \quad \binom{4}{3} \quad \binom{4}{4} \\
\binom{5}{0} \quad \binom{5}{1} \quad \binom{5}{2} \quad \binom{5}{3} \quad \binom{5}{4} \quad \binom{5}{5} \\
\binom{6}{0} \quad \binom{6}{1} \quad \binom{6}{2} \quad \binom{6}{3} \quad \binom{6}{4} \quad \binom{6}{5} \quad \binom{6}{6} \\
\binom{7}{0} \quad \binom{7}{1} \quad \binom{7}{2} \quad \binom{7}{3} \quad \binom{7}{4} \quad \binom{7}{5} \quad \binom{7}{6} \quad \binom{7}{7} \\
\cdots
\end{array}
=
\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 2 \quad 1 \\
1 \quad 3 \quad 3 \quad 1 \\
1 \quad 4 \quad 6 \quad 4 \quad 1 \\
1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6 \quad 1 \\
1 \quad 7 \quad 21 \quad 35 \quad 35 \quad 21 \quad 7 \quad 1 \\
\cdots
\end{array}$$

Aufgabe 3: Modelle der Stochastik untersuchen

a) **Phase 1 (Gruppenarbeit): Bearbeitet** gruppenweise eines der folgenden Arbeitsblätter:

→ **Auslastungsmodell,**
→ **Kugel-Fächer-Modell,**
→ **Geburtstagsparadoxon,**
→ **Warten auf Erfolg.**

b) **Phase 2 (Partnerarbeit): Bearbeite** zusammen mit einem Partner einer anderen Gruppe Deine Aufgabe und die Deines Partners. Wechsele dann zweimal den Partner, so dass Du jede andere Aufgabe kennengelernt hast.

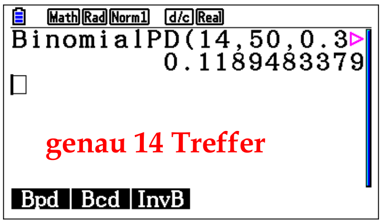
c) **Phase 3 (Gruppenarbeit): Bereitet** in der ursprünglichen Gruppe eine Präsentation Euer Aufgabe vor.

Aufgabe 4: Wahrscheinlichkeiten mit dem GTR berechnen

Betrachten wir nun die binomialverteilte Zufallsgröße X: *„Anzahl der Treffer bei 50 Versuchen jeweils mit der Trefferwahrscheinlichkeit p“*. **Berechne** in dieser 50-stufigen Bernoulli-Kette mit der Trefferwahrscheinlichkeit p die nachfolgenden Wahrscheinlichkeiten, und **gib** ihre Bedeutung **an**.

Zur Berechnung der Wahrscheinlichkeiten nutzen wir den GTR. Im Folgenden werden vier Beispiele für Grundfunktionen mit dem GTR gegeben. Dabei ist die Versuchszahl immer n = 50, die Trefferwahrscheinlichkeit p = 0,3.

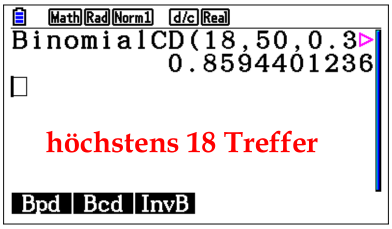
Singuläre Wahrscheinlichkeit P(X = k): P(X = 14) = $\binom{50}{14} \cdot 0{,}3^{14} \cdot 0{,}4^{36}$



Menu 1: OPTN F5 F3 F5 F1 1 4 , 5 0 , 0 . 3) EXE

Erst die Anzahl der Treffer k, dann die Anzahl n der Versuche und zuletzt die Trefferwahrscheinlichkeit p eingeben (immer durch ein Komma getrennt).

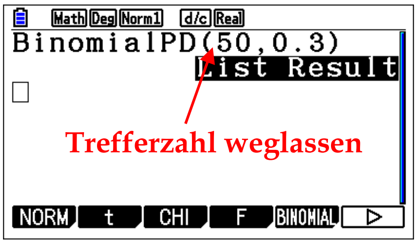
Kumulierte Wahrscheinlichkeit P(X ≤ k): P(X ≤ 18) = P(X = 0)+ …+P(X = 18)

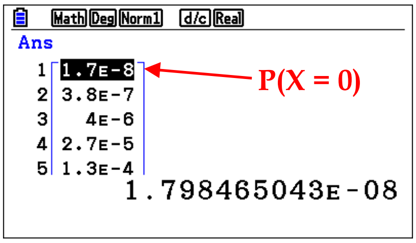


Menu 1: F5 F3 F5 F2 1 8 , 5 0 , 0 . 3) EXE

Erst die Anzahl der (Höchst-)Trefferzahl k, dann die Anzahl n der Versuche und zuletzt die Trefferwahrscheinlichkeit p eingeben (immer durch ein Komma getrennt).

Liste von Wahrscheinlichkeiten P(X = 0), P(X = 1), …, P(X = k) [Gilt durch Ersetzen der Funktion BinomialPD durch BinomialCD analog für eine Liste kumulierter Wahrscheinlichkeit beginnend bei k = 0: P(X ≤ 0), P(X ≤ 1), …, P(X ≤ 50)]

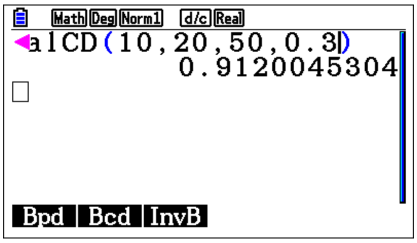




Menu1: OPTN F5 F3 F5 F1 5 0 , 0 . 3) EXE

Trefferzahl weglassen!

Kumulierte Wahrscheinlichkeit P(a ≤ X ≤ b): P(10 ≤ X ≤ 20) = P(X = 10)+ … +P(X = 20)



Menu 1: OPTN F5 F3 F5 F2 1 0 , 2 0 , 5 0 , 0 . 3) EXE

Zuerst untere Grenze, dann obere Grenze der Kumulation, dann die Versuchszahl n und die Trefferwahrscheinlichkeit p angeben.

(1) Trefferwahrscheinlichkeit p = 0,3

Wahrscheinlichkeit	**Wahrscheinlichkeit bei 50 Versuchen …**
$P(X = 14) =$ **0,1189483379**	**… genau 14** Treffer zu erzielen.
$P(X \leq 18) =$ **0,8594401236**	**… höchstens 18** Treffer zu erzielen.
$P(X < 13) =$	
$P(X \geq 17) =$	
$P(X > 21) =$	
$P(10 \leq X \leq 18) =$	

(2) Trefferwahrscheinlichkeit p = 0,6

Wahrscheinlichkeit	**Wahrscheinlichkeit bei 50 Versuchen …**
$P(X = 31) =$	
$P(X \leq 33) =$	
$P(X < 26) =$	
$P(X \geq 27) =$	
$P(X > 31) =$	
$P(23 \leq X \leq 33) =$	

(3) „Komplementäre“ Zufallsgröße Y

Drücke die Wahrscheinlichkeiten in Tabelle (2) zur Zufallsgröße X: *„Anzahl der Treffer mit p = 0,6“* mithilfe der „komplementären“ Zufallsgröße Y: *„Anzahl der Treffer mit p = 0,4“* aus und bestätige die Ergebnisse in Tabelle (2).

Aufgabe 5: Grundaufgaben zur Binomialverteilung

Sei X eine binomialverteilte Zufallsgröße. **Grundaufgaben zur Binomialverteilung** werden durch folgende Tabelle festgelegt, wobei „☑“ eine gegebene Größe und „☒“ die gesuchte Größe repräsentieren:

Grundaufgabe	**1**	**2**	**3**	**4**
Stichprobenumfang n	☑	☑	☑	☒
Trefferwahrscheinlichkeit p	☑	☑	☒	☑
Trefferzahl k	☑	☒	☑	☑
Gesamtwahrscheinlichkeit P	☒	☑	☑	☑

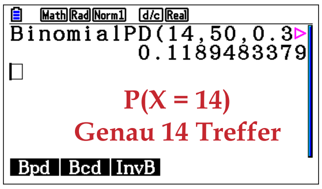
Bevor einige Übungsaufgaben bearbeitet werden sollen, wird zu jeder Grundaufgabe ein Beispiel vorgestellt und dargestellt, wie es mit dem GTR gelöst werden kann. Arbeite die Beispiele durch und löse dann die Anwendungsaufgaben.

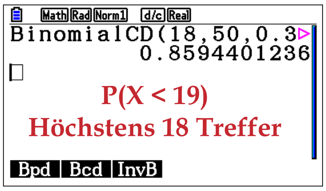
Grundaufgabe 1: Gesamtwahrscheinlichkeit P gesucht

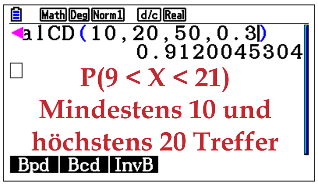
Gegeben: X ist binomialverteilt mit n = 50; p = 0,30 sowie einer bestimmten Trefferzahl k.

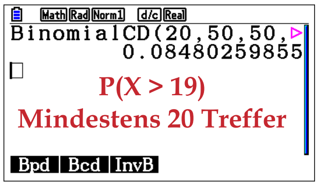
Gesucht: P.

Möchte man z. B. **P(X = 14), P(X < 19), P(9 < X < 21), P(X > 19)** bestimmen, erfolgt dies im MENU 1 (Run – Matrix). Über OPTN und STAT und DIST und BINOMIAL können die obigen Wahrscheinlichkeiten mit den Befehlen BinomialPD und BinomialCD berechnet werden. Diese Befehle kann man auch schneller über SHIFT und 4 (CATALOG) und die Eingabe des Buchstabens B bekommen.







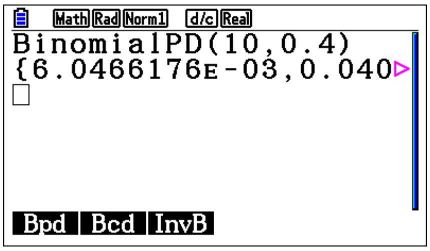


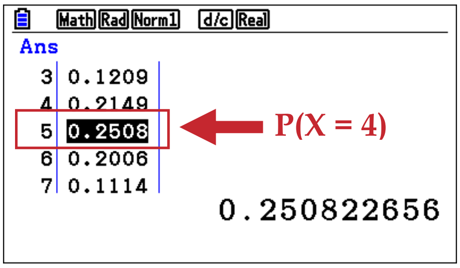
Grundaufgabe 2: Trefferzahl k gesucht

Singuläre Wahrscheinlichkeit

Gegeben: n = 10; p = 0,40 **Gesucht:** Trefferzahl k mit maximaler P(X = k)

Hier kann in **MENU 1** mithilfe der **Listenfunktion** zur singulären Wahrscheinlichkeit gearbeitet werden, indem man nur den Stichprobenumfang n = 10 und die Trefferwahrscheinlichkeit p = 0,4 angibt, aber darauf achten muss, dass die Liste bei 1 startet, was 0 Treffern entspricht.



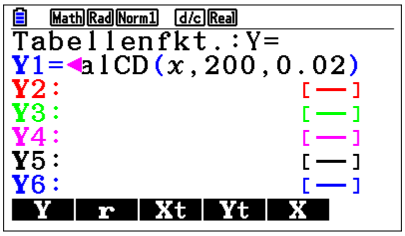


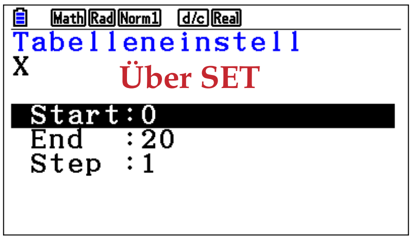
Die größte singuläre Wahrscheinlichkeit beträgt P(X = 4) ≈ 0,2508.

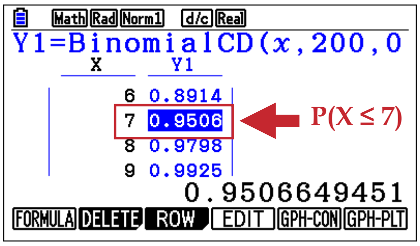
Kumulierte Wahrscheinlichkeit

Gegeben: n = 200, p = 2 %, $P_{n=200;\, p=0{,}02}(X \leq k) \geq 0{,}95$ **Gesucht:** (Mindest-)Trefferzahl k

Die Aufgabe kann über die **Tabellenfunktion (MENU 7)** gelöst werden. Dafür wird eine Funktion f definiert mit f(x) = P(X ≤ x), wobei die Zahl x variiert und die obere Grenze der kumulierten Wahrscheinlichkeit P(X ≤ x) beschreibt. Für x = 7 ist P(X ≤ x) das erste Mal größer als 95 %. Daher gilt: P(X ≤ x) ≥ 0,95 für k ≥ 7.



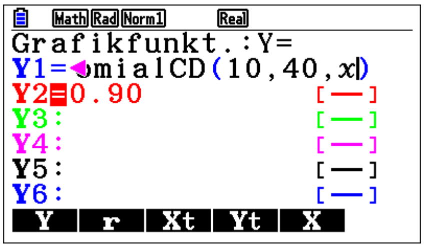


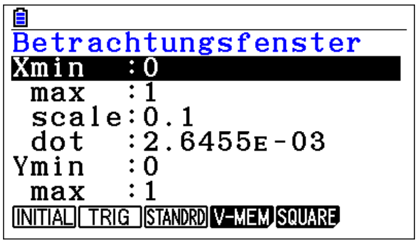


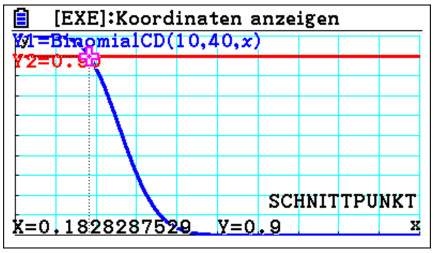
Grundaufgabe 3: Trefferwahrscheinlichkeit p gesucht

Gegeben: n = 40, k = 10; $P_{n=40;\, p}(X \leq k) = 0{,}90$ **Gesucht:** Trefferwahrscheinlichkeit p

Hier kann das **MENU 5 (Graph)** verwendet werden. Es wird dort (unter Y1) eine Funktion f mit $f(x) = P_{n=40;\, x}(X \leq 10)$ definiert sowie (unter Y2) die konstante Funktion g mit der Gleichung g(x) = 0,90, wobei x die variable Trefferwahrscheinlichkeit beschreibt. Die **Schnittstelle** beider Funktionen f und g (erreichbar durch G-Solv und INTSECT) liefert die gesuchte Trefferwahrscheinlichkeit p, denn dort ist $P_{n=40;\, x}(X \leq 10) = 0{,}90$.





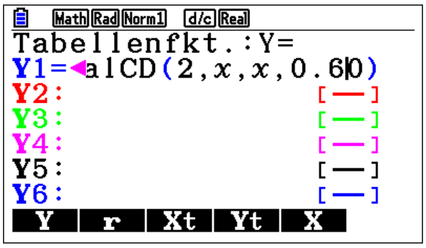


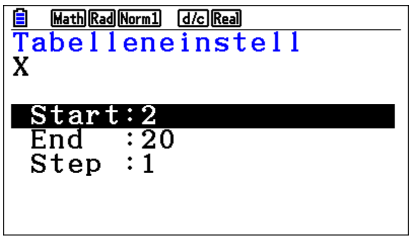
Ist die Trefferwahrscheinlichkeit p ca. 18,28 %, dann ist $P_{n=40;\, p}(X \leq 10) = 0{,}90$.

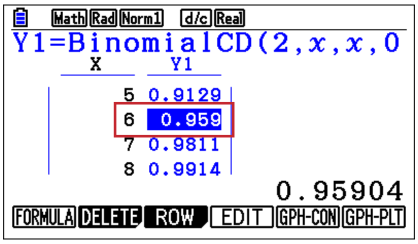
Grundaufgabe 4: Stichprobenumfang n gesucht

Gegeben: k = 2; p = 0,60; $P_{n;p=0{,}80}(X \geq k) \geq 0{,}95$ **Gesucht:** Stichprobenumfang n

Man definiert in MENU 7 (Tabelle) (unter Y1) die Funktion f mit $f(x)=P_{x;p=0{,}80}(X \geq 2)$ und dem variablen Stichprobenumfang x und stellt der Anzeigebereich x über SET entsprechend ein:







Für n ≥ 6 gilt $P_{n;p=0{,}60}(X \geq 2) \geq 0{,}95$.

Nun aber ran an die **Anwendungsaufgaben**:

Grundaufgabe 1: Wahrscheinlichkeit P berechnen (Gegeben: n, k, p)

a) Mit welcher Wahrscheinlichkeit erhält man bei einer Bernoulli-Kette der Länge 20 und der Trefferwahrscheinlichkeit von 30 % genau 5 Treffer?

Wahrscheinlichkeitsgröße X: „*Anzahl der Treffer*", $n = 20$, $k = 5$, $p = 0{,}3 \to P$

b) Ein Tierarzt behandelt 10 kranke Tiere mit einem Medikament, das nach Aussagen des Herstellers in 80% der Anwendungen zur Heilung führt. Mit welcher Wahrscheinlichkeit werden mindestens 9 von 10 Tieren [mindestens 5 und höchstens 8] geheilt?

Wahrscheinlichkeitsgröße X: „*Anzahl geheilter Tiere*", $n = 10$, $p = 80\%$, $k \geq 9$ $[5 \leq k \leq 8] \to P$

Grundaufgabe 2: Trefferzahl k berechnen (Gegeben: n, p, P)

a) Ein Großhändler versorgt 10 Geschäfte, von denen jede Bestellung mit einer Wahrscheinlichkeit von 40% aufgibt. Wie viele Bestellungen laufen mit größter Wahrscheinlichkeit ein?

Wahrscheinlichkeitsgröße X: „*Anzahl der Bestellungen*", $n = 10$, $p = 0{,}4$, P als Höchstwert $\to k$

b) Ein Hersteller von Schrauben behauptet, dass mindestens 98% der Schrauben normgerechte Längen haben. Ein Händler kontrolliert eine Schraubenlieferung mit einer Stichprobe vom Umfang 200 und findet k Schrauben mit nicht normgerechter Länge. Die Rückgabe erfolgt, wenn die Wahrscheinlichkeit für weniger als k nicht normgerechte Schrauben der Stichprobe mindestens 95 % beträgt. Ab welcher Anzahl k sollte er die Lieferung zurückweisen?

X: „*Anzahl nicht normgerechter Schrauben*", $n = 200$, $p = 2\ \%$, $P(X < k) \geq 0{,}95 \to k$

Grundaufgabe 3: Trefferwahrscheinlichkeit p berechnen (Gegeben: n, k, P)

a) Ein Gerät besteht aus 5 Bauteilen, die unabhängig voneinander mit gleicher Funktionswahrscheinlichkeit arbeiten. Fällt ein Bauteil aus, so arbeitet das Gerät nicht mehr. Welche Funktionswahrscheinlichkeit müssen die Bauteile haben, wenn das Gerät mit einer Wahrscheinlichkeit von mehr als 95% funktionieren soll?

Wahrscheinlichkeitsgröße X: „*Anzahl funktionstüchtiger Teile*", $n = 5$, $P \geq 0{,}95$, $k = 5 \to p$

b) Eine Glühlampe, die zufällig der Produktion entnommen wird, leuchtet einwandfrei mit der unbekannten Wahrscheinlichkeit p. Jemand entnimmt zufällig 40 Glühlampen. Mit einer Wahrscheinlichkeit von **mindestens** 90% sollen **mindestens** 38 Glühlampen dieser Stichprobe einwandfrei sein. Wie groß muss die Wahrscheinlichkeit p **mindestens** sein?

Wahrscheinlichkeitsgröße X: „*Anzahl einwandfrei Glühlampen*", $n = 40$, $k = 38$, $P \geq 0{,}9 \to p$

Grundaufgabe 4: Länge n der Bernoulli-Kette berechnen (Gegeben: k, p, P)

a) Auf einer bestimmten Buslinie rechnet man mit 5% Schwarzfahrern. Wie viele Fahrgäste muss ein Kontrolleur mindestens nach ihrem Fahrschein fragen, bis er mit einer Wahrscheinlichkeit von mindestens 90% mindestens einen Schwarzfahrer ertappt hat?

X: „*Anzahl von Schwarzfahrern*", $p = 5\%$, $P \geq 0{,}9$, $k \geq 1 \to n$

b) Ein Schütze trifft das Ziel mit einer Wahrscheinlichkeit von p = 35%. Wie oft muss er auf das Ziel schießen, damit die Wahrscheinlichkeit, das Ziel wenigstens einmal zu treffen, wenigstens 90% beträgt? X: „*Anzahl der Treffer*", $p = 0.35$, $k \geq 1$, $P \geq 0.9 \to n$

c) Ein Zahnarzt weiß, dass die Wahrscheinlichkeit dafür, bei einem Patienten Karies zu diagnostizieren, etwa bei 80% liegt. Wie viele Karteikarten muss man der Patientenkartei zufällig entnehmen, wenn dabei mit einer Wahrscheinlichkeit von mindestens 95% drei oder mehr Patienten mit Kariesbefund sein sollen? X: „*Anzahl von Karies-Patienten*", $p = 80\ \%$, $k \geq 3$, $P \geq 0{,}95 \to n$

Aufgabe 6: Eigenschaften der Binomialverteilung

Ein Kasinobesucher beschwert sich: Die verwendeten sechsseitigen Spielwürfel würden in einem Viertel der Fälle die „1“ zeigen. Die Würfel seien anscheinend gezinkt. Als Test wird einer der Würfel unter gleichen Bedingungen 100-mal geworfen.

Erwartungswert einer Binomialverteilung

a) Wie viele Einser sind zu erwarten? Wie würdest Du den Würfel beurteilen, wenn er 6 Einser zeigt? **Begründe** Deine Antworten.

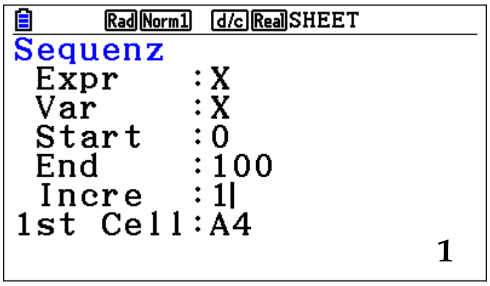
b) **Erläutere**, wie man man allgemein den Erwartungswert $\mu = E(X)$ einer Binomialverteilung (Prognoswert für den Mittelwert der Binomialverteilung) berechnet, wenn X die Anzahl der Treffer in einer n-stufigen Bernoulli-Kette mit der Trefferwahrscheinlichkeit p eschreibt.

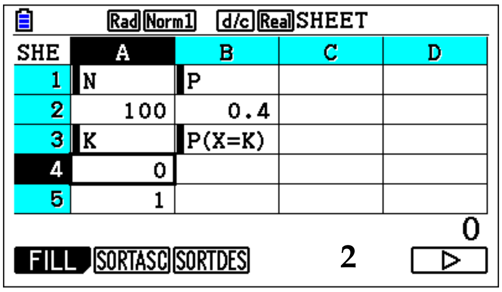
Binomialverteilung in Abhängigkeit von n und p

c) **Verschaffe** Dir einen Überblick über mögliche Testergebnisse, indem Du Histogramme [Streudiagramme] für verschiedene n- und p-Werte zeichnest. Bestimme jeweils den Erwartungswert. **Beschreibe**, was Dir auffällt.

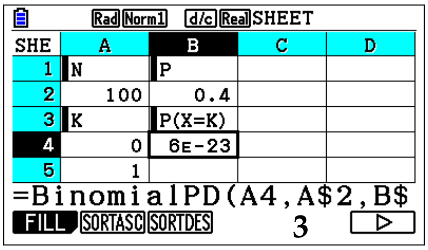
Exemplarisch wird dies für n = 100 und p = 0,4 im MENU 4 (Tabellenkalkulation) demonstriert.

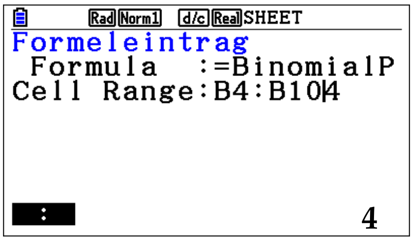
- Nach Eingabe der Spaltenüberschriften (Zum Eintragen von Texten für die Überschriften muss vor dem ersten Buchstaben über $\boxed{\text{ALPHA}}$ und $\boxed{\text{EXP}}$ ein ” stehen) werden Werte für den Stichprobenumfang n und die Trefferwahrscheinlichkeit p eingegeben.
- Man geht auf Feld A4. Über $\boxed{\text{EDIT}}$ und $\boxed{\text{SEQ}}$ kann man die Trefferzahlen 0, 1, …, 60 in die Zellen A4 bis A64 eintragen lassen (Bilder 1 und 2).

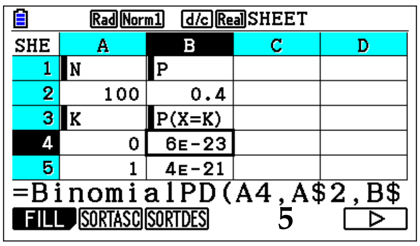




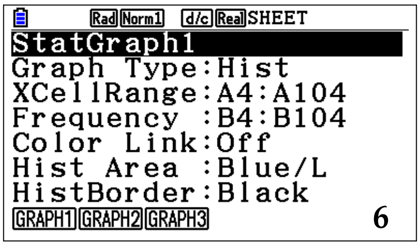
- Nun wird in B4 die Formel für P(X = k) über = BinomialPD(A2,A$2,B$2) eigegeben (Bild 3).
- Über $\boxed{\text{FILL}}$ (vorher $\boxed{\text{EDIT}}$) wird diese Formel in die Zellen B4 bis B104 kopiert (Bilder 4 und 5).



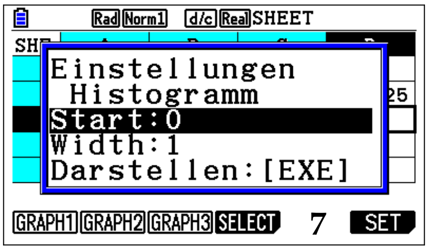




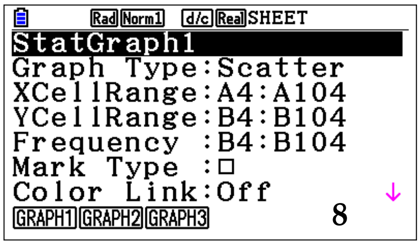
- Nun kann über GRAPH und SET mit folgenden Einstellungen ein **Histogramm** (**Hist,** Bild 6) oder ein Streudiagramm (**Scatter,** Bild 8) erstellt werden (beim Histogramm noch den Startwert 0 und die Schrittweite 1 einstellen, vgl. Bild 7):



6

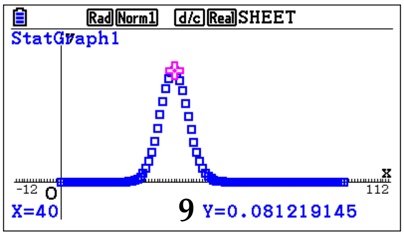


7

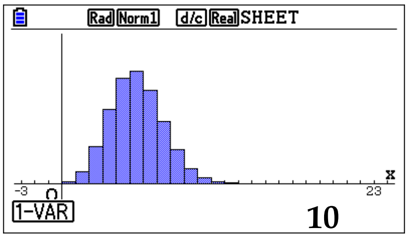


8

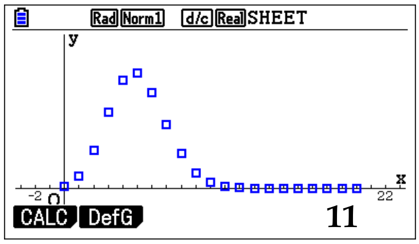
- Über GRAPH1 erhält man das Streudiagramm von Bild 9. Nun lassen sich die Parameter n und p beliebig variieren, wobei n maximal 100 sein kann und p zwischen 0 und 1 liegen muss Ebenso muss über SET der entsprechende Zielbereich (Range) beachtet. Für die Fälle n = 20 und p = 0,25 erhält man als Histogramm Bild 10 und als Streudiagramm Bild 11.



9



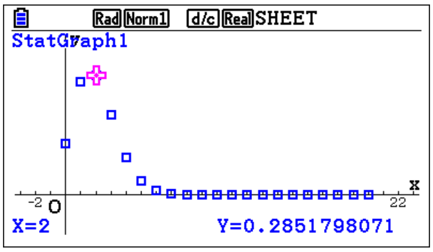
10

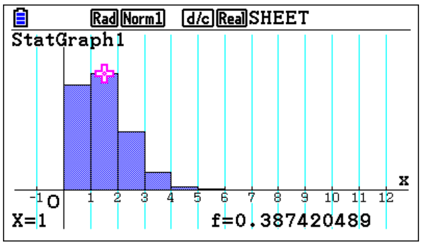


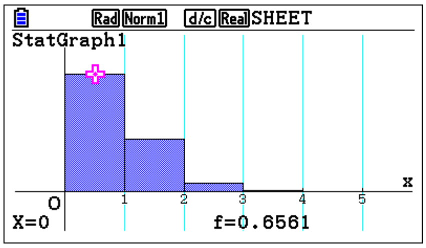
11

Nachfolgend sind einige Fälle (n = 20, n = 10 und n = 4) dargestellt (denke bei einem anderen n an ein neues $\boxed{\text{SET}}$), wobei der Fall n = 20 mithilfe der **Scatter-Funktion** als Streudiagramm dargestellt wird. Die Wahrscheinlichkeitswerte kann man sich mit der **Trace-Funktion** anzeigen lassen.

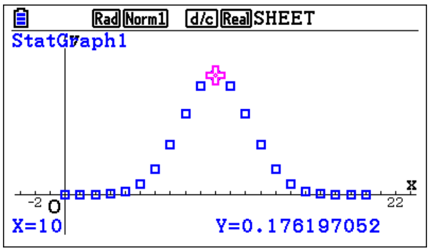
p = 0,1

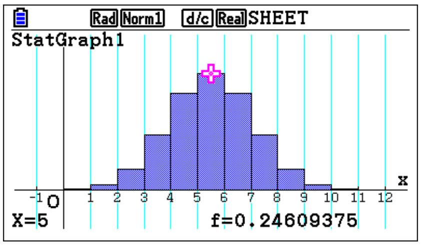


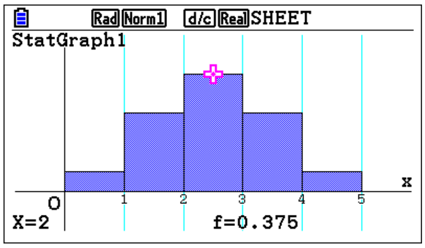




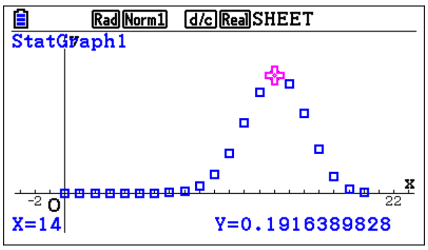
p = 0,5

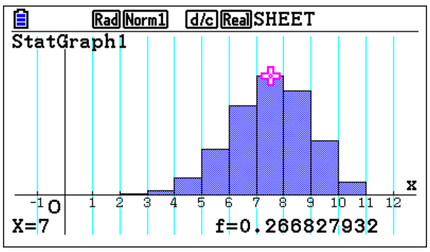


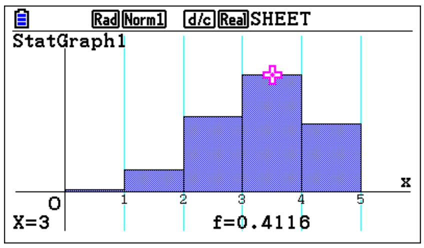




p = 0,7







d) Wichtige Eigenschaften der Binomialverteilung werden in folgendem Text festgehalten. **Übertrage** die Beobachtungen mit passenden Histgrammen in Deinem Heft.

Beobachtungen zur Bionomialverteilung

Die Zufallsgröße X ist die Anzahl der Treffer einer n-stufigen Bernoullikette mit der Trefferwahrscheinlichkeit p.

1) Der Erwartungswert μ der Binomialverteilung berechnet sich durch **$\mu = E(X) = p \cdot n$.**

2) Das Maximum der Binomialverteilung liegt in der Nähe des **Erwartungswertes μ.**

3) Für **p = 0,5** ist die Verteilung symmetrisch.

4) Das Maximum liegt umso weiter rechts, je **größer p ist**.

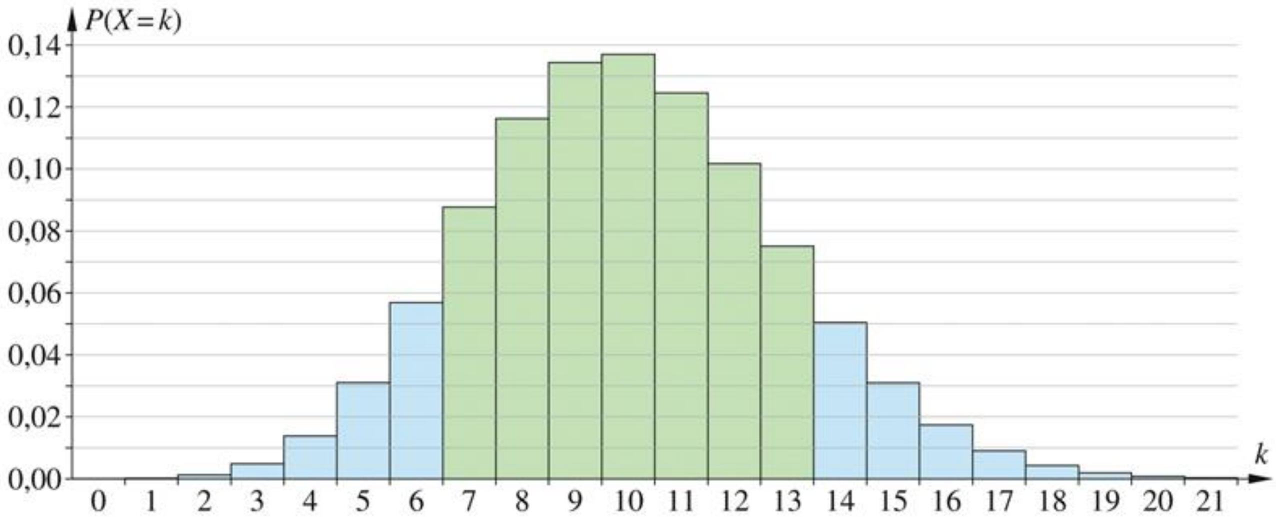
5) Die Binomialverteilung läuft umso flacher, je **größer n ist**.

6) Wird n größer, werden die Histogramme immer symmetrischer zum **Erwartungswert**.

Aufgabe 7: Streuungsmaße der Binomialverteilung

Bei großem n ist das Histogramm sehr flach. Die maximal erreichbare Wahrscheinlichkeit wird mit wachsendem n immer kleiner. Daher ist es recht unwahrscheinlich, dass genau der Erwartungswert eintritt. Vielmehr wird der Ausgang in einem Intervall um den Erwartungswert liegen.

Das folgende Histogramm für den Ausgangstest mit n = 60 und p = $\frac{1}{6}$ zeigt, wie die Versuchsausgänge um den Erwartungswert „streuen" werden.



a) **Bestimme** den Erwartungswert und die Wahrscheinlichkeit, dass zwischen 7 und 13 Einsern gewürfelt werden.

Nun interessiert uns die Breite eines Bereichs um den Erwartungswert (den „Radius der Umgebung um den Erwartungswert"). Wir suchen ein Maß für die Streuung der Zufallsgröße X. Dafür stellen wir folgende Überlegungen an:

- Man berechnet alle Abweichungen der Werte 0, 1, 2, …, n vom Erwartungswert μ:

  $\mathbf{0-\mu, 1-\mu, 2-\mu, \dots, n-\mu}$

- Jede Abweichung $k-\mu$ wird mit der Wahrscheinlichkeit ihres Auftretens $P(X = k)$ gewichtet:

  $\mathbf{(0-\mu)\cdot P(X=0), (1-\mu)\cdot P(X=1), (2-\mu)\cdot P(X=2), \dots, (n-\mu)\cdot P(X=n)}$

- Damit positive und negative Abweichungen sich beim folgenden Aufsummieren nicht gegeneinander aufheben, werden die Abweichungen gleichzeitig noch quadriert:

  $\mathbf{(0-\mu)^2\cdot P(X=0) + (1-\mu)^2\cdot P(X=1) + (2-\mu)^2\cdot P(X=2) + \dots + (n-\mu)^2\cdot P(X=n)}$

- Beim Quadrieren wird leider auch die physikalische Einheit der Zufallsgröße X quadriert. Um diesen Effekt rückgängig zu machen, wird aus der Summe anschließend die Quadratwurzel gezogen. Das so entwickelte Streuungsmaß heißt Standardabweichung von X. Das Symbol ist σ („Sigma“):

  $$\boldsymbol{\sigma}(X) = \sqrt{\mathbf{(0-\mu)^2\cdot P(X=0) + (1-\mu)^2\cdot P(X=1) + \dots + (n-\mu)^2\cdot P(X=n)}}$$

  Das Quadrat der Standardabweichung σ(X) nennt man auch Varianz V(X):

  $$\mathbf{V(X)} = \boldsymbol{\sigma}^2(X) = \mathbf{(0-\mu)^2\cdot P(X=0) + (1-\mu)^2\cdot P(X=1) + \dots + (n-\mu)^2\cdot P(X=n)}$$

Bei einem Stichprobenumfang von n sind zur Berechnung der Varianz n Summanden zu addieren. Für großes n ist dies sehr aufwendig. Zum Glück ist das bei der Binomialverteilung nicht notwendig, wie der folgende Satz zeigt (in Aufgabenteil g) wird der Satz mithilfe der Tabellenkalkulation für verschiedene Werte für n und p nachgewiesen).

Satz zu den Streuungsmaßen einer Binomialverteilung

X sei die Anzahl der Treffer in einer Bernoulli-Kette der Länge n und der Trefferwahrscheinlichkeit p. Ferner sei $q = 1 - p$ die Nichttrefferwahrscheinlichkeit und $\mu = \mathbf{p \cdot n}$ der Erwartungswert von X. Dann gilt für Varianz V(X) und Standardabweichung σ(X):

$$\mathbf{V(X)} = \sigma^2(X) = \mathbf{n\cdot p\cdot q = \mu\cdot q}$$

$$\boldsymbol{\sigma}\mathbf{(X)} = \sqrt{V(X)} = \sqrt{\mathbf{n\cdot p\cdot q}} = \sqrt{\mathbf{\mu\cdot q}}$$

b) **Berechne** nun Standardabweichung σ(X) und Varianz V(X) der 60-stufigen [4-stufige, 20-stufige, 100-stufige] Bernoulli-Kette mit der Trefferwahrscheinlichkeit $p = \frac{1}{6}$ [0,1, 0,25, 0,5, 0,7].

c) **Ergänze** in der Tabellenkalkulation aus Aufgabenteil c) weitere Spalten für $k \cdot P(X = k)$ und $(k-\mu)^2 \cdot P(X = k)$. **Programmiere** darüber hinaus weitere Zellen für die Summe der beiden obigen Produkte von k = 0 bis k = n (Erwartungswert μ, Varianz V), die Quadratwurzel der Varianz (Standardabweichung σ) und die „Nichttrefferwahrscheinlichkeit“ q = 1 – p. Variiere p und n und entdecke einen Zusammenhang von den dargestellten Größen.

4 Sigma-Regeln[1]

Was bedeutet nun die Aussage:

Bei einer Binomialverteilung mit p = 0,4 und n = 125 ist $\sigma = \sqrt{30} \approx 5,48$?

Man kann bei Binomialverteilungen angeben, mit welcher Wahrscheinlichkeit der Versuchsausgang in einer Umgebung des Erwartungswertes liegen sollte, dessen Radius durch ein Vielfaches von σ gegeben ist. Dabei führen wir zunächst den Begriff der **k·σ-Umgebung** ein:

Definition: Eine Umgebung mit dem Radius des k-fachen ($k > 0$) der Standardabweichung σ um den Erwartungswert μ heißt **k·σ-Umgebung (k·σ-Intervall)** und wird durch $[\mu - k\cdot\sigma;\ \mu + k\cdot\sigma]$ gekennzeichnet.

Für Binomialverteilungen gelten die wichtigen Sigma-Regeln. Diese Regeln ermöglichen eine Rechenvereinfachung, da nicht mit kumulierten Wahrscheinlichkeiten gearbeitet werden muss:

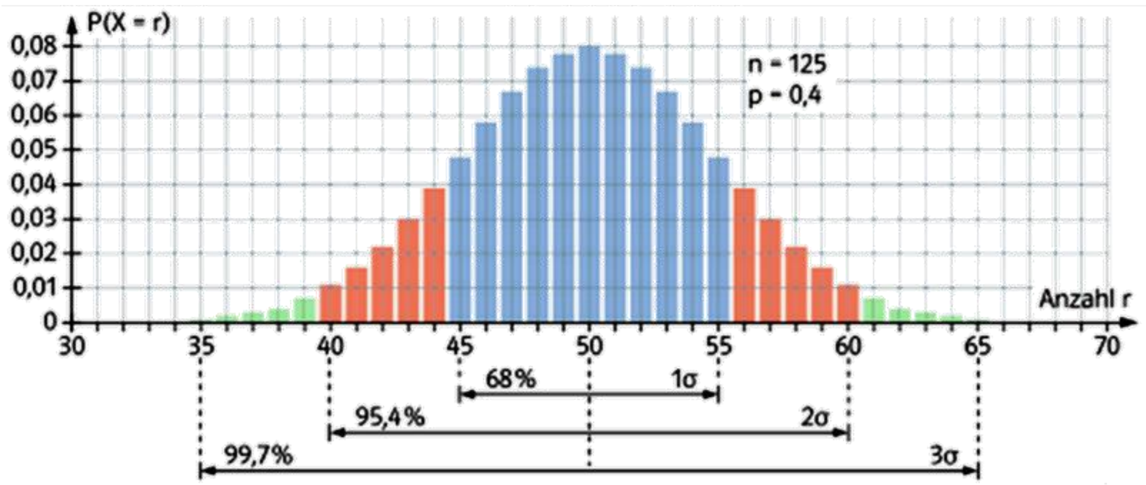
Sigma-Regeln

Für eine binomialverteilte Zufallsgröße X mit den Parametern n und p, dem dazugehörigen Erwartungswert $\mu = n \cdot p$ und der entsprechenden Standardabweichung $\sigma = \sqrt{n \cdot p \cdot (1-p)}$ erhält man folgende Näherungen, falls **$\sigma > 3$ (Laplace-Bedingung)** gilt:

1σ-Regel: $P(\mu - \sigma \le X \le \mu + \sigma) \approx 68{,}3\ \%$
2σ-Regel: $P(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx 95{,}4\ \%$
3σ-Regel: $P(\mu - 3\sigma \le X \le \mu + 3\sigma) \approx 99{,}7\ \%$

Für „glatte" Wahrscheinlichkeiten gelten die folgenden Sigma-Regeln:

$P(\mu - 0{,}67\sigma \le X \le \mu + 0{,}67\sigma) \approx 50\ \%$
$P(\mu - 0{,}84\sigma \le X \le \mu + 0{,}84\sigma) \approx 60\ \%$
$P(\mu - 1{,}04\sigma \le X \le \mu + 1{,}04\sigma) \approx 70\ \%$
$P(\mu - 1{,}15\sigma \le X \le \mu + 1{,}15\sigma) \approx 75\ \%$
$P(\mu - 1{,}28\sigma \le X \le \mu + 1{,}28\sigma) \approx 80\ \%$
$P(\mu - 1{,}64\sigma \le X \le \mu + 1{,}64\sigma) \approx 90\ \%$
$P(\mu - 1{,}96\sigma \le X \le \mu + 1{,}96\sigma) \approx 95\ \%$
$P(\mu - 2{,}33\sigma \le X \le \mu + 2{,}33\sigma) \approx 98\ \%$
$P(\mu - 2{,}58\sigma \le X \le \mu + 2{,}58\sigma) \approx 99\ \%$
$P(\mu - 3{,}29\sigma \le X \le \mu + 3{,}29\sigma) \approx 99{,}9\ \%$



[1] Fakultativ im GK

Mithilfe der Standardabweichung σ können Intervalle um den Erwartungswert $\mu = E(X)$ eines Zufallsexperiments gelegt werden, in denen Trefferzahlen mit bestimmten Wahrscheinlichkeiten liegen. Dabei gilt folgender **Merksatz**:

> Liegt die Trefferwahrscheinlichkeit bei einem Bernoulli-Experiment **außerhalb der 2σ-Umgebung (2σ-Intervall)**[2], spricht man von einer **signifikanten Abweichung** vom Erwartungswert. Das bedeutet: der Ausgang ist so ungewöhnlich, dass vermutlich ein anderes p als das angenommene zugrunde lag.

Beispiel: Würde man bei 125 Versuchen im obigen Beispiel weniger als 40 ($50 - 2\sigma \approx 39{,}04$) oder mehr als 60 Treffer ($50 + 2\sigma \approx 60{,}96$) erhalten, müsste man von einer anderen Trefferwahrscheinlichkeit ausgehen. Es gilt $P(40 \le X \le 60) \approx 94{,}5\ \%$ (runde zur „sicheren“ Seite nach innen ab), so dass in 5,5 % der Fälle weniger als 40 und mehr als 60 Treffer eintreten.

Übungsaufgaben zu den Sigma-Regeln

a) Ein angenommener nicht „gezinkter“ Hexaeder-Würfel wird 100 Mal gewürfelt. Es kommt 10 [bzw. 25] Mal die „1“.

   Untersuche, ob eine signifikante Abweichung vorlag und wie hoch die Trefferzahl höchstens [bzw. mindestens] sein darf, dass sie mit einer Wahrscheinlichkeit von mindestens 99 % in einer geeigneten Umgebung um den Erwartungswert liegt.

b) Auf der Kirmes wirbt eine Losbude mit dem Versprechen: Jedes dritte Los gewinnt! Johannes ist misstrauisch und kauft 60 Lose, um die Aussage zu testen. Darunter sind nur 12 Gewinnlose.

   Beurteile die Werbung mithilfe der 2σ-Regel.

c) Johanna plant, eine Münze 100 Mal zu werfen. Sie möchte das Intervall wissen, in das die Zahl X ihrer Kopfwürfe fällt, mit mindestens 95 % Sicherheit voraussagen.

   Untersuche, welches Intervall sie angeben sollte.

d) Der Wochenspiegel ist eine kostenlose Werbezeitung mit einer Auflage von 5000 Exemplaren. In 80 % der Haushalte wandert sie ungeöffnet in die Mülltonne.

   Bestimme die Wahrscheinlichkeit, dass höchstens 980 bzw. mindestens 1100 Exemplare gelesen werden. **Schätze** mit der 2σ-Regel, in welchem Intervall die Anzahl gelesener Exemplare mit ca. 95,4 %-iger Wahrscheinlichkeit liegt.

e) **Bearbeite** die Aufgaben a) bis d) ohne Sigma-Regeln nur mit kumulierten Wahrscheinlichkeiten.

[2] Außerhalb der 3σ-Umgebung (3σ-Intervall) spricht man von einer hoch signifikanten Abweichung vom Erwartungswert.

5 Testen von Hypothesen mithilfe der Binomialverteilung

Aufgabe 1: Rotes oder grünes Gummibärchen?

Wählt in Eurer Tischgruppe eine Person aus, die mit verbundenen Augen sagen soll, ob es sich um ein rotes oder ein grünes Gummibärchen handelt. Dieser Versuch wird 40-mal wiederholt.

a) **Notiert** das Ergebnis.

b) **Vermutet**, mit welcher Wahrscheinlichkeit die Person die Gummibärchen richtig erkennt. Könnte die Person die Erfolgsquote auch durch reines Raten erreicht haben? Unter welchen Bedingungen handelt es sich bei dem Versuch um ein Bernoulli-Experiment?

c) **Bestimme** Erwartungswert, Varianz und Standardabweichung.

Aufgabe 2: Hypothesenbildung beim Geschmackstest[3]

Wir nehmen an, dass die beiden obigen Experimente eine Bernoulli-Kette der Länge 40 beschreiben. Im Folgenden betrachten wir den Geschmackstest der Gummibärchen.

a) **Berechne** näherungsweise, wie wahrscheinlich es ist, mindestens 23, 27 bzw. 30 der 40 Gummibärchen richtig einer Farbe zuzuordnen.

b) **Entscheide** begründet: Kann man ab einer bestimmten Anzahl k richtig zugeordneter Farben das Raten ausschließen? Wenn ja, ab wann?

Was ist im obigen Kontext eine Hypothese?

Das Vorgehen soll nun verallgemeinert werden und der Begriff der **Hypothese** eingeführt werden. Wir wissen nicht, wie hoch die Wahrscheinlichkeit wirklich ist, die Gummibärchen richtig einer Farbe zuzuordnen. Daher können wir nur sagen, wie wahrscheinlich es ist, die beobachtete Anzahl x_{beob} an richtig zugeordneten Gummibärchen oder mehr („mindestens x_{beob}") zu erreichen, wenn der Schüler nur rät. Man stellt daher folgende **Hypothesen** auf:

Nullhypothese H_0	**Alternativhypothese H_A**
Die Person rät nur.	Die Person ist besser als Raten.

Welche mathematische Testarchitektur erhält man bzw. wie erfolgt die Modellbildung?

Wandelt man diese Hypothesen in **statistische** Formulierungen um, erhält man in der obigen Situation des Geschmackstests folgende mathematische Beschreibungen:

Nullhypothese H_0: $p = 0{,}5$	**Alternativhypothese H_A:** $p > 0{,}5$
Die Person ordnet alle Gummibärchen mit einer konstanten Erfolgswahrscheinlichkeit $p = 0{,}5$	Dem Prozess der Farbzuordnung eines Gummibärchens könnte eine Wahrscheinlichkeit zu-

[3] Die Beispielaufgabe „Gummibärchen" ist angelehnt an die Fortbildungsreihe des DZLM: „STOCHASTIK konkret 2014" von Rolf Biehler und Janina Oesterhaus

der richtigen Farbe zu und die einzelnen Versuche sind stochastisch unabhängig voneinander. Dadurch sind die Bedingungen zur Anwendung der Binomialverteilung erfüllt und in dem Modell kann für die Zufallsgröße X „Anzahl der richtig zugeordneten Gummibärchen“ die Versuchsanzahl n = 40 verwendet werden.	grunde liegen, die p > 0,5 beträgt. Eine genauere Festlegung ist im Zusammenhang schwierig (da z. B. Erfahrung, Übung, etc. eine konstante Wahrscheinlichkeit nicht sicher erscheinen lässt). Deshalb wird die Alternativhypothese erstmal nicht weiter festgelegt, sondern bei „besser als raten“ belassen.

Wann wird eine Hypothese angenommen, wann wird sie abgelehnt?

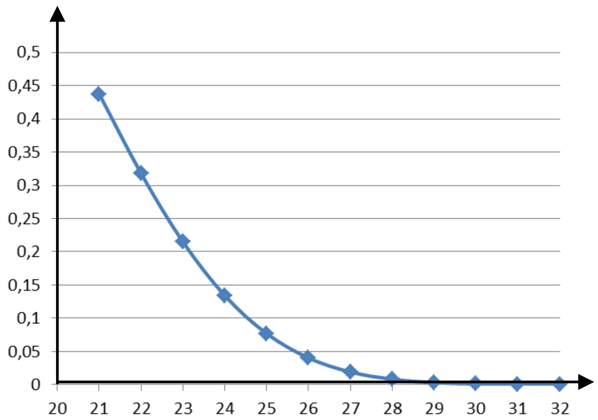
Wenn die Anzahl an richtig zugeordneten Gummibärchenfarben vom Erwartungswert 20 nach „links“ abweicht (also geringer als „raten“ ausfällt), kann davon ausgegangen werden, dass die Person nur rät (und das noch nicht mal mit besonders viel Glück!). Daher ist in unserem Fall nur der Bereich rechts vom Erwartungswert 20 interessant. Nun muss ein **Annahme**- und der **Verwerfungsbereich** der Nullhypothese bestimmt werden. Ein gängiger Annahmebereich liegt bei 5%. Wir sagen dann, dass das **Signifikanzniveau α bei 5%** liegt.[4] Die Wahrscheinlichkeit, die Hypothese zu verwerfen, obwohl sie zutrifft, nennt man **Irrtumswahrscheinlichkeit**. Sie ist die Wahrscheinlichkeit des Ablehnungsbereichs und beträgt daher höchsten 5 %. Man sagt, dass die **Abweichung** (von H_0) **statistisch signifikant auf dem 5%-Niveau** ist. Die maximale Irrtumswahrscheinlichkeit entspricht daher dem Signifikanzniveau α.

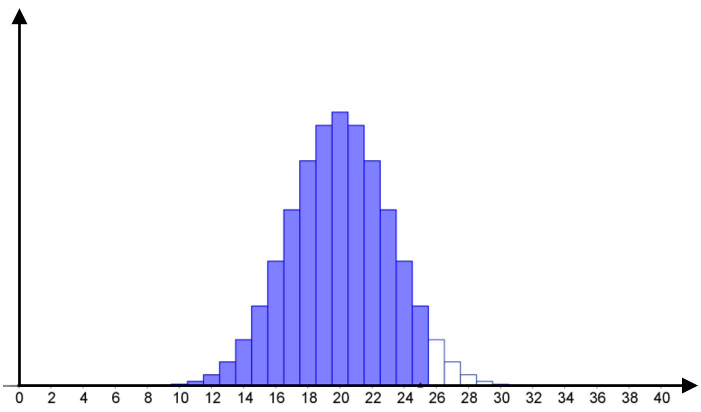
c) **Bestimme** die Wahrscheinlichkeiten in der Tabelle für p = 0,5 und n = 40 und gib den Annahme- und Verwerfungsbereich der Nullhypothese an.

k	21	22	23	24	25	26	27	28	29	30	31
P(x = k)											
P(X ≤ k)											
P(X ≥ k)											

Beurteile aufgrund der obigen Tabelle, für welche der drei Ergebnisse $x_{beob} = 23, 27, 30$ die Nullhypothese angenommen werden kann und wann sie verworfen werden muss.

d) **Ergänze** in den beiden Diagrammen die Achsenbeschriftung bzw. -markierung und interpretiere sie im obigen Kontext.





e) **Entscheide** begründend, welche der folgenden Aussagen wahr bzw. falsch sind.

[4] Hier liegt die (willkürlich festgelegte, aber übliche) Annahme zugrunde, dass der Verwerfungsbereich insgesamt 5% betragen soll. Für einseitige Tests bedeutet das ein 5%-Intervall, für zweiseitige Tests 2,5%-Intervall an jeder Seite.

Für die Aussagen (1) bis (5) wird angenommen, dass ein Testergebnis im Verwerfungsbereich auf einem Signifikanzniveau von 5 % liegt.

Aussage	Wahr	Falsch	Begründung
1. Es ist eindeutig bewiesen, dass die Nullhypothese falsch ist.	☐	☐	
2. Die Wahrscheinlichkeit des Zutreffens der Nullhypothese ist gefunden worden.	☐	☐	
3. Es ist eindeutig bewiesen, mit der die Alternativhypothese wahr ist.	☐	☐	
4. Man kann nun die Wahrscheinlichkeit ableiten, dass die Alternativhypothese richtig ist.	☐	☐	
5. Man geht nun davon aus, dass die Nullhypothese falsch ist, ist sich aber bewusst, dass man bei diesem Verfahren eine wahre Nullhypothese mit einer Wahrscheinlichkeit von maximal 5% fälschlich verwirft.	☐	☐	

Für die Aussagen (6) bis (10) wird angenommen, dass ein Testergebnis im Annahmebereich auf einem Signifikanzniveau von 5 % liegt.

Aussage	Wahr	Falsch	Begründung
6. Es ist eindeutig bewiesen, dass die Nullhypothese richtig ist	☐	☐	
7. Man kann nun die Wahrscheinlichkeit ableiten, dass die Nullhypothese richtig ist.	☐	☐	
8. Es ist eindeutig bewiesen, dass die Alternativhypothese falsch ist.	☐	☐	
9. Die Wahrscheinlichkeit des Zutreffens der Alternativhypothese ist gefunden worden.	☐	☐	
10. Man geht nun davon aus, dass die Nullhypothese richtig ist, ist sich aber bewusst, dass man bei diesem Verfahren in dem Fall, in dem die Nullhypothese falsch ist, diese mit einer Wahrscheinlichkeit von 5% dennoch fälschlich annimmt	☐	☐	

Aus den bisherigen Überlegungen lässt sich folgende Reihung für einen **rechtsseitigen Signifikanztest** einer **Nullhypothese** H_0: $p = p_0$ und seiner **Alternative** H_A: $p > p_0$ ableiten:

1. Man legt den Stichprobenumfang n und das Signifikanzniveau α (z. B. $\alpha = 5\%$) fest.
2. Als Testgröße X verwendet man die Trefferzahl für die Parameter n und p_0.
3. Man bestimmt den Annahmebereich der Nullhypothese. Dazu sucht man aus der Tabelle der kumulierten Wahrscheinlichkeiten von X die kleinste Zahl b heraus, bei der $1 - \alpha = 95\%$ überschritten wird. Die Irrtumswahrscheinlichkeit beträgt dann höchstens bei $\alpha = 5\%$.
4. Man führt die Stichprobe vom Umfang n durch. H_0 wird angenommen, wenn das Stichprobenergebnis im Annahmebereich liegt, sonst wird H_0 verworfen und H_A angenommen.

Was versteht man unter einem Fehler 1. und 2. Art?

Beim Testen mit Binomialverteilungen wird die Nullhypothese akzeptiert oder verworfen. Dabei können Fehlentscheidungen vorkommen. Wenn die Nullhypothese verworfen wird, obwohl sie richtig ist, spricht man von einem **Fehler 1. Art (α-Fehler)**. Wenn sie akzeptiert wird, obwohl sie falsch ist, spricht man von einem **Fehler 2. Art (β-Fehler)**. Folgende Tabelle stellt dies dar.

		Zustand der Wirklichkeit	
		H_0 ist **wahr**	H_0 ist **falsch**
H_0-Hypothese wird …	**abgelehnt**	**Fehler 1. Art / α-Fehler**	**richtige Entscheidung**
	beibehalten	**richtige Entscheidung**	**Fehler 2. Art / β-Fehler**

Beispiele:

- Im obigen Fall der Gummibärchen tritt ein **Fehler erster Art** auf, wenn aufgrund eines außergewöhnlich guten Ratens (z.B. 28 Treffer) fälschlicherweise angenommen wird, dass die Person die Farbe der Gummibärchen mit einer Wahrscheinlichkeit von mehr als 50 % schmecken kann.
- Ein **Fehler zweiter Art** tritt auf, wenn der Proband z. B. 21 Treffer erreicht (was im Annahmebereich von H_0 liegt), aber in Wirklichkeit doch besser als Raten ist.

Wie berechnet man α- und β-Fehler?

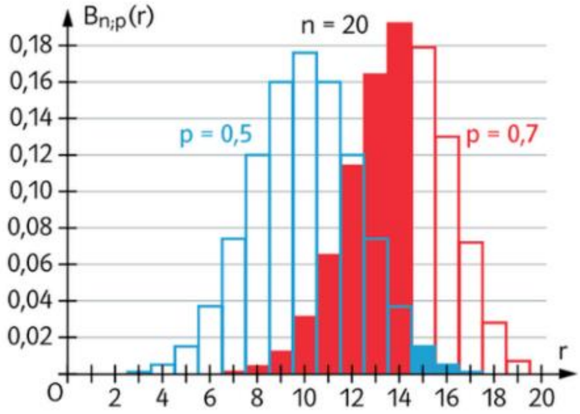
Der **Fehler erster Art** ist inhaltlich das gleiche wie die Irrtumswahrscheinlichkeit und somit durch das Signifikanzniveau festgelegt. Er beträgt $1 - P(X \leq 25) = \mathbf{P(X \geq 26) \approx 0,0403}$ (was natürlich kleiner als 0,05 sein muss).

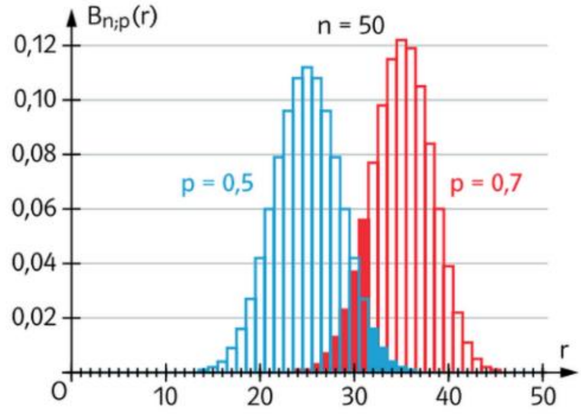
f) **Ermittle** den Fehler der ersten Art für einen höheren bzw. niedrigeren Stichprobenumfang n. **Beschreibe** Deine Beobachtungen.

Der **Fehler zweiter Art** hängt von der tatsächlichen Wahrscheinlichkeit p ab, mit der die Versuchsperson die Gummibärchenfarbe erkennen würde. Diese ist in aller Regel unbekannt. Deshalb muss man sich einen Überblick verschaffen und in Abhängigkeit einer angenommenen (hier: besseren) Trefferquote den Fehler bestimmen. Sei X_p die Anzahl der Treffer bei einer angenommenen variablen Trefferquote p und einem festem Stichprobenumfang n. Es gilt für den β-Fehler $P(X_p \leq 25)$. Für $p = 0,6$ gilt: $\mathbf{P(X_{0,6} \leq 25) \approx 0,6825}$.

g) **Bestimme** für n = 40 [n = 20; n = 100] den Fehler der 2. Art für unterschiedliche Wahrscheinlichkeiten p > 0,5. **Beschreibe** Deine Beobachtungen.

h) **Beschreibe**, wie sich der Fehler 1. Art und 2. Art verändern, wenn der Annahmebereich vergrößert wird. zusammenhängen. **Erläutere**, wie sich die beiden Fehler verändern, wenn der Stichprobenumfang vergrößert wird. Nutze dafür die folgenden Abbildungen von Binomialverteilungen.





Wie und wann helfen die Sigmaregeln?

i) Die **Sigmaregeln** geben die Wahrscheinlichkeit dafür an, dass die Trefferzahl bei einer Bernoullikette der Länge n bestimmte symmetrische Abweichungen (der Größe $a \cdot \sigma$)[5] vom Erwartungswert μ nicht überschreitet. Im Falle, dass die Standardabweichung $\sigma > 3$ ist, gilt:

a	**P_a („zweiseitig“) $= P(\mu - a \cdot \sigma < X < \mu + a \cdot \sigma)$**	**P_a („einseitig“) $= 1 - P(X < \mu - a \cdot \sigma) = P(X \geq \mu - a \cdot \sigma)$ („links“) $= 1 - P(X > \mu + a \cdot \sigma) = P(X \leq \mu + a \cdot \sigma)$ („rechts“)**
3,29	99,9%	99,95%
3	99,7%	99,85%
2,58	99%	99,5%
2,33	98%	99%
2	95,4%	97,7%
1,96	95%	97,5%
1,64	90%	95%
1,28	80%	90%
1,15	75%	87,5%
1,04	70%	85%
1	68,3%	84,2%
0,84	60%	80%
0,67	50%	75%

Ermittle mithilfe der Sigmaregeln den Annahme- und Verwerfungsbereich eines rechtsseitigen Tests auf einem Signifikanzniveau von 5% [10%; 20%; 1%] für n = 40 und p = 0,5.

---

[5] Die Wahrscheinlichkeiten sind für $\sigma > 3$ praktisch unabhängig von der Versuchslänge n und der Trefferwahrscheinlichkeit p. Für die Normalverteilungen sind die Gleichungen exakt.

Aufgabe 3: Tombola

Für eine Tombola auf dem Schulfest wird eine große Menge an Losen gekauft. Die Herstellerfirma teilt mit, dass der Anteil der Nieten 65% beträgt. Durch eine Stichprobe von 80 Losen will sich der Mathe-Leistungskurs nun von der Richtigkeit der Angaben überzeugen. Gesucht ist eine Entscheidungsregel für eine Nullhypothese auf einem Signifikanzniveau von 1%.

a) **Beschreibe** die **Testarchitektur** der obigen Situation.

b) **Gib** die **Nullhypothese H_0** für den obigen Sachkontext **an**.

c) Die **Gegenhypothese H_A** ist keine Festlegung auf einen bestimmten Anteil der Nieten. Sie wird nur angenommen, wenn das Ergebnis der Stichprobe mit einer Wahrscheinlichkeit von 99% darauf schließen lässt, dass die Nullhypothese nicht zutrifft. **Formuliere** H_A mathematisch.

d) Die Aufgabe ist so konstruiert, dass der Mathematik-LK herausfinden will, ob die Nieten tatsächlich mit einem Anteil von 65% in den Losen erhalten sind. Das bedeutet, dass es von Interesse ist, ob nicht doch *mehr oder weniger* Nieten vorhanden sind. Somit befinden wir uns bei einem **zweiseitigen Signifikanztest**.

(1) **Bestimme** die Werte für die Zufallsgröße X, den Umfang der Stichprobe und den Erwartungswert μ und die Standardabweichung σ. **Notiere** dazu nochmals die Wahrscheinlichkeit p sowie das vorgegebene Signifikanzniveau α.

(2) **Bestimme** mit dem GTR für die singulären und kumulierten Wahrscheinlichkeiten der folgenden Tabelle. [Zur Erinnerung: „BinomialPD(80, 0,65)“ bzw. „BinomialCD(80, 0,65)“ zeigen die Liste der singulären bzw. kumulierten Wahrscheinlichkeiten an.]

	…	39	40	41	42	43	…	62	63	64	65	…
$P(x = k)$	…						…					…
$P(X \leq k)$	…						…					…
$P(X \geq k)$	…						…					…

e) Die Wahrscheinlichkeit dafür, dass (bei angenommener H_0) das Ergebnis im Verwerfungsbereich landet, soll der Aufgabe zufolge höchstens 1% betragen. Daraus ergeben sich für die linke und rechte Seite höchstens 0,5%.

(1) **Gib** in der Liste der kumulierten Wahrscheinlichkeiten die kleinste Zahl a **an**, bei der 0,005 überschritten wird.

(2) **Bestimme** in dieser Liste die kleinste Zahl b, bei der 0,995 überschritten wird.

(3) **Formuliere** den Annahme- sowie den Verwerfungsbereich der Nullhypothese sowie eine Entscheidungsregel (im ganzen Satz).

f) **Für Experten: Untersuche**, was es bedeutet hätte, wenn der Mathe-LK das Signifikanzniveau auf 10% festgelegt hätte. Begründe, welche Reaktion gegenüber der Herstellerfirma möglich wäre, wenn das Ergebnis bei einem 1% Signifikanzniveau im Vergleich zu einem 10% Signifikanzniveau im Verwerfungsbereich landen würde.

g) **Für Experten:** Der Mathe-LK führt zunächst die Stichprobenüberprüfung aus und wählt dann das Signifikanzniveau. **Bewerte** das Vorgehen.

Was sind Fehler 1. Art und 2. Art beim beidseitigen Testen?

h) Der **Fehler erster Art (α-Fehler)**, also die **„Irrtumswahrscheinlichkeit“**, ergibt sich aus dem Signifikanzniveau und gibt die Wahrscheinlichkeit dafür an, dass H_0 verworfen wird, obwohl es wahr ist.

(1) **Formuliere** den Fehler erster Art im Anwendungszusammenhang.
(2) **Berechne** den Fehler erster Art.

i) Der **Fehler zweiter Art** gibt die Wahrscheinlichkeit dafür an, dass H_0 akzeptiert wird, obwohl es falsch ist.

(1) **Begründe**, dass der Fehler zweiter Art nicht berechnet werden kann und **formuliere** die Bedeutung des Fehlers zweiter Art im Sachzusammenhang.
(2) Nimm an, dass tatsächlich 70% Nieten im Lossatz vorhanden sind und **bestimme** den Fehler zweiter Art.
(3) **Erläutere** ohne Rechnung, wie sich im Vergleich zu b) der Fehler zweiter Art verändert, wenn tatsächlich 80% Nieten im Lossatz vorhanden sind.

Wie bestimmt man die Entscheidungsregel mit den Sigmaregeln?

j) Alternativ zu c) und d) kann der Annahme- und Verwerfungsbereich auch mit den **Sigmaregeln** bestimmt werden.

(1) **Berechne** die Standardabweichung σ.
(2) **Bestimme** aus der in Aufgabe 3 i) von Kapitel 1 angegebenen Tabelle das zum 1%-Signifikanzniveau zugehörige a.
(3) **Berechne** damit den Annahme- und den Verwerfungsbereich.

Aufgabe 4: Verspätungen im Nahverkehr[6]

Die Bürgerinitiative „NRW für ÖPNV“ behauptet, dass die Verkehrsmittel im Nahverkehr in mindestens 40% der Fahrten eine merkliche Verspätung (von mehr als 5 Minuten) hätten. Der VRR bestreitet das. Es soll eine Untersuchung von 100 zufällig ausgewählten Fahrten erfolgen und der Test soll auf einem Signifikanzniveau von 5% erfolgen.

a) **Beschreibe** die „**Testarchitektur**“ und formuliere **Nullhypothese** H_0 und **Alternativhypothese** H_A mathematisch.

b) **Bestimme** die Werte für die Zufallsgröße X, den Umfang der Stichprobe und den Erwartungswert μ sowie die Standardabweichung σ.

c) **Bestimme** mit dem GTR die **kumulierte Binomialverteilung** für alle $k \in \{0, \dots, 100\}$.

d) **Formuliere** den **Annahme**- sowie den **Verwerfungsbereich** der Nullhypothese sowie eine **Entscheidungsregel** (im ganzen Satz).

[6] Die Beispielaufgabe „Verspätungen“ ist an das Skript der Kompetenzteams-NRW-Federführungsgruppe der Bezirksregierung Köln (2011) angelehnt.

e) **Formuliere** und berechne den **Fehler erster Art** im Anwendungszusammenhang (unter der Voraussetzung, dass H_0 fälschlicherweise verworfen wurde).

f) **Formuliere** und **berechne** den **Fehler zweiter Art** unter der Annahme, dass die tatsächliche Wahrscheinlichkeit für Verspätungen 30% beträgt (und H_0 fälschlicherweise angenommen wurde). **Bewerte** den Fehler zweiter Art aus der Perspektive der Bürgerbewegung und des VRR.

g) **Berechne** mit Hilfe der Sigma-Regeln den Annahme- und den Verwerfungsbereich.

h) **Für Experten:** Die Bürgerbewegung behauptet, dass es in *mindestens* 40% aller Fahrten zu einer Verspätung kommt. Die Bahn hingegen behauptet, dass dies in *höchstens* 40% der Fahrten der Fall ist. Ein (rechtsseitiger) Signifikanztest auf 5%-Signifikanzniveau ergibt einen Annahmebereich von A = [0;48] mit einem Fehler erster Art von ca. 4% und einem Fehler zweiter Art von 10%. **Bewerte** diesen Testansatz aus der Perspektive der Bürgerbewegung.

Aufgabe 5: Münzwurf (Grundlagen)

Es werden fünfzigmal zwei Münzen gleichzeitig geworfen. Die Zufallsvariable X zählt, wie oft beide Münzen „Zahl" zeigen.

a) **Begründe**, warum dies eine Bernoulli-Kette ist und geben Sie die Parameter an.

b) **Berechne** die Wahrscheinlichkeit, dass
   (1) genau zwölfmal beide Münzen „Zahl" zeigen,
   (2) höchstens zwölfmal beide Münzen „Zahl" zeigen,
   (3) mindestens zwölfmal beide Münzen „Zahl" zeigen
   (4) mindestens fünfzehnmal und höchstens dreißigmal beide Münzen „Zahl" zeigen.

c) **Berechne** den Erwartungswert von X. Wie kann man diesen Wert interpretieren?

d) **Berechne** die Standardabweichung von X.

e) **Gib** ein Intervall **an**, in dem mit einer Wahrscheinlichkeit von ca. 95% die Anzahl der Würfe liegt, bei denen beide Münzen „Zahl" zeigen.

Aufgabe 6: Kirschkerne (Grundlagen)

In einem Unternehmen werden Sauerkirschen maschinell entsteint und dann in Gläser abgefüllt. 2% der fertigen Kirschen haben trotzdem noch ihren Kern.

a) Herr Becker backt einen Kirschkuchen. Dafür nimmt er 100 dieser Kirschen. **Untersuche**, wie groß die Wahrscheinlichkeit ist, dass sich in dem Kuchen mindestens ein Kirschkern befindet.

b) **Bestimme**, wie viele Kirschkerne (im Mittel auf lange Sicht) in einem solchen Kuchen zu erwarten sind.

c) **Untersuche**, wie viele Kirschen Herr Becker für seinen Kuchen höchstens nehmen dürfte, damit er mit mindestens 80% Wahrscheinlichkeit keinen Kern darin hat.

d) **Ermittle**, wie groß die Wahrscheinlichkeit, dass eine Kirsche noch ihren Kern hat, sein darf, damit sich mit mindestens 80% Wahrscheinlichkeit in einem Kirschkuchen mit 100 Kirschen kein Kirschkern befindet.

Aufgabe 7: Brot auf die Marmeladenseite

Katja möchte testen, ob ein Marmeladenbrot mit gleicher Wahrscheinlichkeit auf die unbestrichene oder die mit Marmelade bestrichene Seite fällt. Dazu hat sie 25 Toastbrote auf einer Seite mit jeweils gleich viel Marmelade bestrichen und möchte einen zweiseitigen Signifikanztest mit dem Signifikanzniveau 5% durchführen.

a) **Gib** die Nullhypothese und die Alternative **an**.

b) **Untersuche**, bei welchen Stichprobenergebnissen Katja davon ausgeht, dass beide Seiten gleich wahrscheinlich sind.

c) **Berechne** die Wahrscheinlichkeit für den Fehler erster Art.

d) **Ermittle** die Wahrscheinlichkeit für den Fehler zweiter Art, wenn das Brot tatsächlich mit der Wahrscheinlichkeit 0,7 auf die Marmeladenseite fällt.

Aufgabe 8: Umfrage zur Stadthalle

Der Oberbürgermeister einer Stadt behauptet, dass 75% der Bürger für den Bau einer neuen Stadthalle sind. Die Redaktion der Lokalzeitung glaubt, dass es weniger sind. Sie möchte dazu einen Signifikanztest in Form einer Umfrage unter 100 Bürgern der Stadt durchführen. Das Signifikanzniveau soll 5% betragen.

a) **Begründe**, ob ein linksseitiger, ein rechtsseitiger oder ein zweiseitiger Test durchgeführt werden soll und **gib** die Nullhypothese, die Alternative und den Annahmebereich des Tests **an**.

b) **Beurteile**, bei welchen Ergebnissen die Redaktion die Schlagzeile „Weniger Befürworter als OB behauptet“ drucken kann.

c) **Berechne** die Wahrscheinlichkeit für den Fehler erster Art.

d) **Bestimme** die Wahrscheinlichkeit für den Fehler zweiter Art, falls nur 60% der Bürger für den Bau sind.

Überblicksraster für Signifikanztests

Aufgabe: Erstelle in Deinem Heft jeweils Musterbeispiele für einen rechts- und linksseitigen Signifikanztest wie im folgenden Beispiel zum zweiseitigen Signifikanztest geschehen.

Musterbeispiel für zweiseitigen Signifikanztest auf binomialverteilten Zufallsgrößen

Betrachtet wird eine beliebige Beispielaufgabe zu einem zweiseitigen Signifikanztest

1. Aufstellen der Hypothesen:

Nullhypothese im Zusammenhang: H_0: p = 0,3

Gegenhypothese im Zusammenhang: H_A: p ≠ 0,3

2. Aufstellen der Variablen:

n = 50 p = 0,3 $\alpha = 0{,}05 \Rightarrow \frac{\alpha}{2} = 0{,}025$

Die Prüfvariable *X* ist binomialverteilt mit n = 50 und p = 0,3.

3. Bestimmung von Annahme und Verwerfungsbereich:

Stellt man mit dem GTR die kumulierte Verteilung dar, so ergibt sich:

	…	6	7	8	**9**	10	…	21	**22**	23	24	…
P(X ≤ k)	**…**	0,0024	0,0073	0,0183	**0,0402**	0,0789	…	0,9749	**0,9877**	0,9944	0,9976	**…**

Als erstes ist das **kleinste a** gesucht mit $\mathbf{P(X \leq a) \geq \frac{\alpha}{2} = 0,025}$. Dies ist für **a = 9** der Fall.
Dann sucht man das **kleinste b** mit $\mathbf{P(X \leq b) \geq 0,975}$. Dies ist für **b = 22** der Fall.

Dann ist der **Verwerfungsbereich** $V = \{0,1,2, \dots, 8\} \cup \{23,24,25, \dots, 50\}$ und der **Annahmebereich** $A = \{10,11,12, \dots ,19,20,21,22\}$.

4. Bewertung des Tests

- Da das Testergebnis 9 im Annahmebereich liegt, wird man die Nullhypothese annehmen.
- Da das Testergebnis 6 im Verwerfungsbereich liegt, wird man die Nullhypothese ablehnen.

5. Fehlerberechnung

Fehler 1. Art:
$\alpha = P(X \leq 8) + P(X \geq 23) = P(X \leq 8) + 1 - P(X \leq 22) = 0{,}0183 + 1 - 0{,}9877 = 0{,}0306$

Fehler 2. Art: p_{neu} sei die tatsächliche Trefferwahrscheinlichkeit. Dann gilt:

$\beta = P_{p_{neu}}(9 \leq X \leq 22) = P_{p_{neu}}(X \leq 22) - P_{p_{neu}}(X \leq 8)$

6 Kontrollaufgaben

Kompetenzraster zur Vorbereitung auf die Klausur

Hilfsmittelfrei

Ich kann …	Aufgabe	sicher	ziemlich sicher	unsicher	sehr unsicher
Kenngrößen von Häufigkeitsverteilungen berechnen.	1				
Kenngrößen von Wahrscheinlichkeitsverteilungen berechnen.	1				
Wahrscheinlichkeiten bei Urnenversuchen mit Zurücklegen berechnen.	2				
bedingte Wahrscheinlichkeiten mittels Vierfeldertafel berechnen.	3				
Bernoulli-Terme bestimmten Ereignissen zuordnen.	4				
Histogramme einer bestimmten Binomialverteilung zuordnen.	5				
Wahrscheinlichkeiten anhand von Histogrammen bestimmen.	5				
Zusammenhänge von komplementären Zufallsgrößen herstellen.	5				
Kenngrößen von Binomialverteilungen (n und p gegeben) berechnen.	6				
Kenngrößen von Binomialverteilungen (n und V gegeben) berechnen.	6				
Kenngrößen an einem Histogramm zur Binomialverteilung ablesen.	6				

Unter Nutzung des GTR

Ich kann …	Aufgabe	sicher	ziemlich sicher	unsicher	sehr unsicher
erläutern, ob das Modell der Binomialverteilung anwendbar ist.	9a,10b				
Grundaufgabe 1 einer Binomialverteilung lösen (Gesucht: P).	9b,c,10b,c				
Grundaufgabe 2 einer Binomialverteilung lösen (Gesucht: k).	9e,h,10c				
Grundaufgabe 2 einer Binomialverteilung lösen (Gesucht: p).	9d				
Grundaufgabe 2 einer Binomialverteilung lösen (Gesucht: n).	9f				
Erwartungswert berechnen.	9b,10b				
Baumdiagramme zu einem Sachverhalt darstellen.	10a				
Wahrscheinlichkeiten mittels Baumdiagramm bestimmen.	10a				
bedingte Wahrscheinlichkeiten berechnen.	9g, 10a				
Sigma-Regeln auf binomialverteilte Zufallsgröße anwenden.	9i				
Fehler 1. Art und 2. Art im Sachkontext beschreiben.	10c				
Fehler 1. Art mittels Entscheidungsregel berechnen	10c				
Kenngrößen einer Häufigkeitsverteilung berechnen.	10d				
Rückschlüsse aus Kenngrößen einer Häufigkeitsverteilung ziehen.	10d				
kombinatorische Aufgaben zum Lotto 6 aus 49 lösen.	11				

Aufgaben ohne Zuhilfenahme von Hilfsmitteln

Aufgabe 1: Kenngrößen von Häufigkeits- und Wahrscheinlichkeitsverteilungen

a) In den Schulaufgaben in Mathematik und Englisch hat sich in einer Klasse die folgende Notenverteilung ergeben:

Note	1	2	3	4	5	6
Mathematik	6	4	2	3	4	6
Englisch	2	4	6	6	4	2

Welche Antworten sind richtig? **Begründe**.

(1) Das arithmetische Mittel von M und E ist gleich.
(2) Die Standardabweichung von M und E ist gleich.
(3) Die Standardabweichung von M ist größer als die von E.

b) Eine Zufallsgröße X habe die folgende Wahrscheinlichkeitsverteilung.

x_i	0	1	2
$P(X = x_i)$	0,32	0,36	0,32

(1) **Begründe**, warum es sich um keine Binomialverteilung handelt.

(2) **Bestimme** den Erwartungswert μ, die Varianz V(X) und die Standardabweichung σ(X) der Wahrscheinlichkeitsverteilung von X.

Aufgabe 2: Urnenversuche

In einer Urne befinden sich 6 rote und 4 blaue Kugeln.

a) Es wird dreimal eine Kugel mit Zurücklegen gezogen.

Bestimme die Wahrscheinlichkeit für das Ereignis E_1 „Unter den gezogenen Kugeln ist höchstens eine blaue Kugel".

b) Es wird zehnmal eine Kugel mit Zurücklegen gezogen. Als Ereignis werde betrachtet E_2: „Unter den gezogenen Kugeln sind genau k blaue Kugeln ($0 \le k \le 4$).

Gib eine Formel für die Wahrscheinlichkeit des Ereignisses E_2 **an**.

Aufgabe 3: Wahrscheinlichkeiten

In einer Reisegruppe sind 37,5 % männliche Reisende (M), von diesen sind 80 % im Alter von 60 und mehr (60+). Insgesamt sind 70 % der Reisende im Alter 60+.

a) **Bestimme** den Anteil der weiblichen Reisenden im Alter 60+ in der gesamten Reisegruppe.

[Tipp: Vierfeldertafel]

b) Insgesamt gibt es 10 mehr weibliche als männliche Reisende in der Gruppe.

Ermittle die Personenzahl der gesamten Reisegruppe.

Aufgabe 4: Bernoulli-Formel

a) **Entscheide begründend**, welche der folgenden Terme zu dem angegebenen Ereignis passen.

Die Wahrscheinlichkeit, dass von 20 zufällig ausgewählten Personen genau drei Männer sind beträgt (wir nehmen an, dass es genauso viele Männer wie Frauen gibt) …

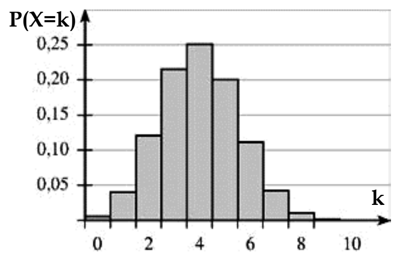
(1) $\binom{20}{3} \cdot 0{,}5^3 \cdot 0{,}5^{17}$ (2) $0{,}5^3 \cdot 0{,}5^{17}$ (3) $\binom{20}{17} \cdot 0{,}5^3 \cdot 0{,}5^{17}$ (4) $\binom{20}{3} \cdot 0{,}5^{20}$ (5) $\frac{20 \cdot 19 \cdot 18}{3!} \cdot 0{,}5^3 \cdot 0{,}5^{17}$

b) Ein Schnellrestaurant veranstaltet ein Gewinnspiel und schenkt jedem Kunden ein Los. Die Wahrscheinlichkeit für einen Sofortgewinn liegt bei $\frac{1}{5}$. Ordne den folgenden Ereignissen den richtigen Term zur Berechnung der Wahrscheinlichkeit zu: A: Unter 10 Losen sind keine Sofortgewinne; B: Unter 10 Losen sind genau 4 Sofortgewinne; C: Unter 10 Losen ist mindestens ein Sofortgewinn.

$T_1 = \binom{10}{4} \cdot \left(\frac{1}{5}\right)^4 \cdot \left(\frac{4}{5}\right)^6$ $T_2 = 1 - \left(\frac{4}{5}\right)^{10}$ $T_3 = \left(\frac{4}{5}\right)^{10}$ $T_4 = \binom{10}{4} \cdot \left(\frac{1}{4}\right)^5 \cdot \left(\frac{3}{4}\right)^5$ $T_5 = 1 - \left(\frac{1}{5}\right)^{10}$ $T_6 = \left(\frac{1}{5}\right)^{10}$

Aufgabe 5: Binomialverteilung

Die Zufallsvariablen X und Y sind binomialverteilt. Für die Variable X ist n = 10 und p = 0,4. Für die zweite Variable Y ist n = 10 und p = 0,6. Die Wahrscheinlichkeitsverteilung einer der beiden Variablen ist in der folgenden Abbildung dargestellt.



a) **Begründe,** dass die Wahrscheinlichkeitsverteilung der Zufallsgröße X abgebildet ist, und **gib** einen Wert für $P(X \neq 5)$ **an.**

b) **Begründe**, warum $P(X = k) = P(Y = 10 - k)$ für alle natürlichen Zahlen $0 \leq k \leq n$ gilt, und **bestimme** $P(Y = 6)$.

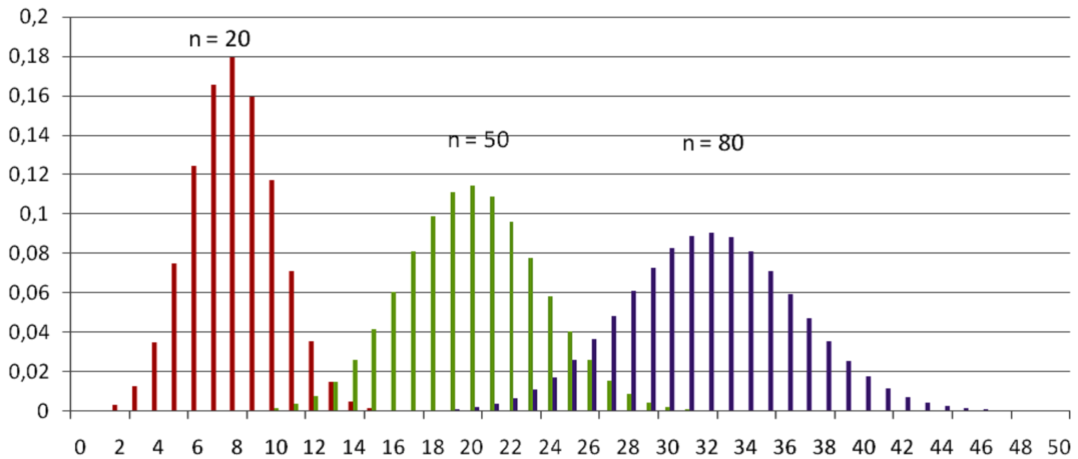
Aufgabe 6: Erwartungswert, Varianz, Standardabweichung (Binomialverteilung)

a) **Bestimme** für eine binomialverteilte Zufallsvariable X mit n = 100 und p = 0,2 den Erwartungswert μ und die Standardabweichung σ.

b) Die Zufallsgröße Y sei binomialverteilt mit dem Parameter n = 100. Y habe die Varianz V(Y) = 9.

Berechne die Trefferwahrscheinlichkeit p und mögliche Erwartungswerte μ.

c) Die Grafik zeigt die Säulendiagramme dreier Binomialverteilungen. Bei allen ist p = 0,4. Welche Verteilung hat die größte, welche die kleinste Standardabweichung? **Begründe** Deine Antwort.



Aufgaben unter Berücksichtigung von Hilfsmitteln

Aufgabe 7: Multiple-Choice-Test

Bei einem Multiple-Choice-Test werden n Aufgaben mit jeweils vier Antworten zur Auswahl gestellt. Von den vier Antworten ist genau eine Antwort richtig. Die Fragen befassen sich mit einem begrenzten Stoffgebiet.

a) Der Bewerber A hat bezüglich des Stoffgebietes einen Kenntnisstand von 70 %.

   Erkläre, unter welche Annahmen man den Test als n-stufige Bernoulli-Kette behandeln kann.

b) **Berechne** die Wahrscheinlichkeit, dass …

   (1) er mindestens 60 % der Fragen richtig beantwortet, wenn der Test aus 20 Fragen besteht.
   (2) sich sein Testergebnis um höchstens 5 Antworten vom erwarteten unterscheidet.

c) **Bestimme** die Wahrscheinlichkeit, dass ein gut trainierter Bewerber B (Kenntnisstand 80 %) …

   (1) von 10 Fragen, die ersten 8 richtig beantwortet und die letzten beiden falsch.
   (2) von 10 Fragen 8 richtig beantwortet.

d) Ein dritter Bewerber C beantwortet mit einer Wahrscheinlichkeit von **mindestens** 90 % von 20 Fragen **mindestens** 15 richtig.

   Ermittle, welchen Kenntnisstand der Bewerbers C **mindestens** besitzt.

e) **Untersuche**, wie viele der 20 Fragen ein vierter Bewerber D mit einem Kenntnisstand von 90 % mit größter Wahrscheinlichkeit beantwortet.

f) **Bestimme** die Zahl der Fragen, die ein Bewerber E mit einem Wissensstand von 75 % gestellt bekommen muss, damit er mindestens 10 Fragen mit einer Wahrscheinlichkeit von mindestens 95 % beantwortet.

g) Es stellt sich heraus, dass 80 % der Bewerber den Test bestehen. Die Bewerber wurden danach alle noch weiter beobachtet. Von den Bewerbern, die den Test bestanden haben, bewährten sich auch 80 % weiterhin. Von denen, die den Test nicht bestanden haben, bewährten sich immerhin noch 20 % in der Folgezeit.

   Berechne den Anteil der Bewerber, die sich bewährt haben. [Tipp: Vierfeldertafel]

h) Unser anfangs genannter Bewerber A kommt in die nähere Auswahl. Er hat dabei noch einen weiteren Test mit nun 60 Fragen zu beantworten, wobei jede Frage von der bewertenden Personalchefin mit einem Schwierigkeitsgrad von 0,6 eingestuft wird. Leider kann sie nur 75 % der verbleibenden Bewerber einstellen, und zwar die mit den besten Testergebnissen bei diesem letzten Test.

   Untersuche, wie viele richtige Antworten sie verlangen muss. (LK)

i) Die Daten des dritten Tests werden ausgewertet. Die Personalchefin möchte nun wissen, wie groß die Abweichung vom durchschnittlich zu erwartenden Ergebnis maximal sein darf, damit keine signifikante Abweichung vom Erwartungswert vorliegt und der angenommene Schwierigkeitsgrad bestätigt wird.

   Beantworte die Fragestellung der Personalchefin mithilfe der 2σ-Regel. (LK)

ABITUR 2017 Aufgabe 8: Glutenunverträglichkeit[7]

In Deutschland liegt bei 1 % der Bevölkerung eine Glutenunverträglichkeit vor. Die betroffenen Personen reagieren auf den Verzehr von bestimmten Getreidesorten mit körperlichen Beschwerden. Ob eine Glutenunverträglichkeit vorliegt oder nicht, kann mithilfe eines Schnelltests diagnostiziert werden. Zeigt das Ergebnis dieses Tests die Glutenunverträglichkeit an, so bezeichnet man es als positiv.

a) Liegt bei einer Person eine Glutenunverträglichkeit vor, so ist das Testergebnis mit einer Wahrscheinlichkeit von 98 % positiv. Liegt bei einer Person keine Glutenunverträglichkeit vor, so beträgt die Wahrscheinlichkeit dafür, dass das Testergebnis dennoch positiv ist, 4 %. Bei einer Person, die aus der Bevölkerung Deutschlands zufällig ausgewählt wurde, wird der Test durchgeführt.

(1) *Erstellen Sie zu dem beschriebenen Sachzusammenhang ein beschriftetes Baumdiagramm.*

(2) *Ermitteln Sie für folgende Ereignisse jeweils die Wahrscheinlichkeit:*

A: „Bei der Person liegt eine Glutenunverträglichkeit vor und das Testergebnis ist positiv.“
B: „Das Testergebnis ist negativ.“

(3) *Ermitteln Sie die Wahrscheinlichkeit dafür, dass eine Glutenunverträglichkeit vorliegt, wenn das Testergebnis positiv ist.*

b) Im Rahmen einer Studie sollen aus der Bevölkerung Deutschlands 20000 Personen zufällig ausgewählt werden. Die Zufallsgröße X ist binomialverteilt ($p = 0{,}01$) und gibt die Anzahl der ausgewählten Personen an, bei denen eine Glutenunverträglichkeit vorliegt.

(1) *Erläutern Sie, warum das Modell der binomialverteilten Zufallsgröße zur Berechnung von Wahrscheinlichkeiten hier geeignet ist.* (LK)

(2) *Bestimmen Sie für folgende Ereignisse jeweils die Wahrscheinlichkeit:*

E_1: Bei genau 190 Personen liegt eine Glutenunverträglichkeit vor.
E_2: Bei mehr als 19800 Personen liegt **keine** Glutenunverträglichkeit vor.
E_3: Mindestens 240, aber höchstens 2400 Personen besitzen eine Glutenunverträglichkeit.

(3) *Bestimmen Sie die Wahrscheinlichkeit dafür, dass die Anzahl der ausgewählten Personen, bei denen eine Glutenunverträglichkeit vorliegt, um mehr als 10 % vom Erwartungswert von X abweicht.*

c) Der Test wird mithilfe eines Teststreifens durchgeführt, auf dem eine Substanz als Indikator aufgebracht ist. Ist die Indikatormenge auf einem Teststreifen zu gering, so ist dieser unbrauchbar. Der Hersteller der Teststreifen verfolgt das Ziel, dass höchstens 10 % der hergestellten Teststreifen unbrauchbar sind, und führt deshalb regelmäßig eine Qualitätskontrolle durch. Dazu wird der laufenden Produktion eine Stichprobe von 100 Teststreifen entnommen. Nur wenn sich darunter mindestens 16 unbrauchbare Teststreifen befinden, entscheidet man sich dafür, das Herstellungsverfahren zu verbessern.

(1) *Beschreiben Sie, welche Fehlentscheidungen bei dieser Qualitätskontrolle auftreten können.* (LK)

(2) *Bestimmen Sie, wie hoch die Fehlerwahrscheinlichkeit höchstens ist, dass der Hersteller sich aufgrund seiner Entscheidungsregel irrtümlich um eine Verbesserung des Herstellungsverfahrens bemüht.* [Kontrolllösung: 3,99 %]

---

[7] Modifiziert nach GK und LK HT B5 Zentralabitur NRW 2017

(3) Der Hersteller entschließt sich, die Kontrolle künftig mit einer größeren Stichprobe von 200 Teststreifen durchzuführen. Die Wahrscheinlichkeit für eine unnötige Verbesserung des Herstellungsverfahrens soll durch diese Änderung nur noch ein Drittel der ursprünglichen Wahrscheinlichkeit für eine unnötige Verbesserung des Verfahrens betragen.

Ermitteln Sie, wie groß die Anzahl unbrauchbarer Teststreifen, ab der man sich dafür entscheidet, das Herstellungsverfahren zu verbessern, nun mindestens sein muss. (LK)

(4) Durch einen Maschinendefekt sind statt 10 % nun 18 % der Teststreifen unbrauchbar.

Bestimmen Sie die Wahrscheinlichkeit, dass dieser Defekt bei Beibehaltung der Entscheidungsregel fälschlicherweise nicht bemerkt wird.

d) Im Rahmen der Qualitätskontrolle wird u. a. die Indikatormenge auf den einzelnen Teststreifen gemessen. Tabelle 1 zeigt die absoluten Häufigkeiten der aufgetretenen Mengen bei einer Stichprobe von 100 Teststreifen.

Indikatormenge in mg	15	16	17	18	19	20
Anzahl der Teststreifen	4	9	10	48	18	11

(1) *Bestimmen Sie für diese Häufigkeitsverteilung das arithmetische Mittel und die Standardabweichung.*

(2) Bei einer früheren Qualitätskontrolle lag das arithmetische Mittel bei 18 mg und die Standardabweichung befand sich bei 4,3 mg.

Erläutern Sie unter Berücksichtigung Ihrer Ergebnisse aus (1), welche Rückschlüsse sich aus diesen Kenngrößen auf die Qualitätsentwicklung des Produktionsverfahrens ziehen lassen.

Aufgabe 9: Lotto 6 aus 49

Berechne die Wahrscheinlichkeit im Lotto „6 aus 49“, …

a) 6 Richtige,

b) 4 Richtige (LK),

c) keine richtige Zahl zu tippen (LK).

Lösungen

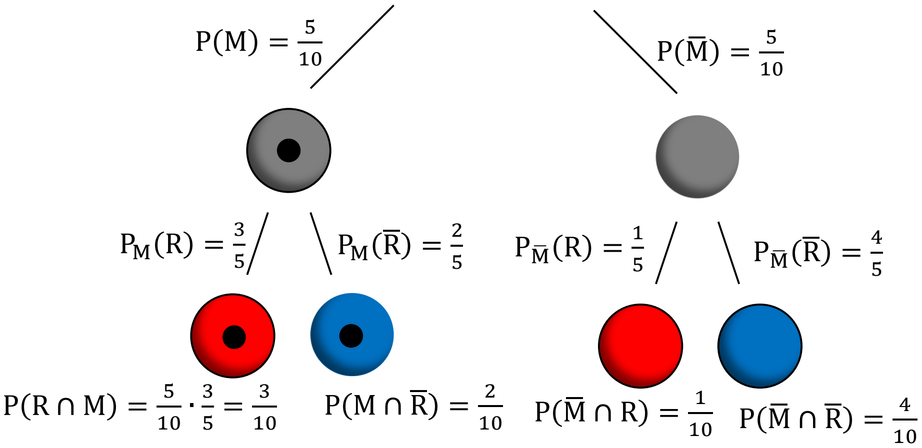
1 Noch fit? – Stochastisches Grundwissen aus der E-Phase

Aufgabe 3

a)

$P(M) = \frac{5}{10} = 50\,\%,\ P(\overline{M}) = \frac{5}{10} = 50\,\%, P_R(M) = \frac{3}{4} = 75\,\%, P_R(\overline{M}) = \frac{1}{4} = 25\,\%, P_{\overline{R}}(M) = \frac{2}{6} = \frac{1}{3},$
$P_{\overline{R}}(\overline{M}) = \frac{4}{6} = \frac{2}{3}, P_M(R) = \frac{3}{5} = 60\,\%, P_M(\overline{R}) = \frac{2}{5} = 40\,\%, P_{\overline{M}}(R) = \frac{1}{5} = 20\,\%,\ P_{\overline{M}}(\overline{R}) = \frac{4}{5} = 80\,\%,$
$P(R) = \frac{4}{10} = 40\,\%,\ P(\overline{R}) = \frac{6}{10} = 60\,\%,\ P(R \cap M) = \frac{3}{10} = 30\,\%,\ P(R \cap \overline{M}) = \frac{1}{10} = 10\,\%,$
$P(\overline{R} \cap M) = \frac{2}{10} = 20\,\%,\ P(\overline{R} \cap \overline{M}) = \frac{4}{10} = 40\,\%$

b)



c)

Sensitivität: $P_{\text{Infiziert}}(+) = \frac{P(\text{Infiziert} \cap +)}{P(\text{Infiziert})} = \frac{0{,}000999}{0{,}001} = 99{,}9\%$
Spezifität: $P_{\text{Nicht infiziert}}(-) = \frac{P(\text{Nicht infiziert} \cap -)}{P(\text{Nicht infiziert})} = \frac{0{,}997002}{0{,}999} = 99{,}8\%$

Aufgabe 4

a)

	A (Stürmer)	$\overline{A}$ (kein Stürmer)	
Startelf (B)	0,40	0,40	**0,8**
Einwechselspieler ($\overline{B}$)	0,15	0,25 · 0,2 = 0,05	0,2
	0,55	0,45	

b)

Der Torschütze ist …
(1) Stürmer der Startelf: $P(A \cap B) = 0{,}40$
(2) weder Stürmer noch Einwechselspieler: $P(\overline{A} \cap B) = 0{,}40$

(3) Stürmer oder Einwechselspieler: : $P(A \cup \overline{B}) = 0{,}40 + 0{,}15 + 0{,}05 = 0{,}60$
(4) kein Stürmer: $P(\overline{A}) = 0{,}45$
(5) entweder Stürmer oder Einwechselspieler: $P\big((\overline{A} \cap B) \cup (A \cap \overline{B})\big) = 0{,}40 + 0{,}05 = 0{,}45$
(6) nicht gleichzeitig Einwechselspieler und Nichtstürmer: $P\big(\overline{\overline{A} \cap \overline{B}}\big) = 1 - 0{,}05 = 0{,}95$.

c)

Wahrscheinlichkeit, dass …
(1) unter den Stürmern ein Einwechselspieler getroffen hat: $P_A(\overline{B}) = \frac{P(A \cap \overline{B})}{P(A)} = \frac{0{,}15}{0{,}55} = \frac{3}{11} \approx 27{,}3\ \%$.
(2) unter den Einwechselspielern ein Nichtstürmer getroffen hat: $P_{\overline{B}}(\overline{A}) = \frac{P(\overline{A} \cap \overline{B})}{P(\overline{B})} = \frac{0{,}05}{0{,}20} = \frac{1}{4} \approx 25\ \%$
(3) unter den Startspielern ein Stürmer getroffen hat: $P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{0{,}40}{0{,}80} = 0{,}50 = 50\ \%$.
(4) unter den Nichtstürmern ein Startspieler getroffen hat: $P_{\overline{A}}(B) = \frac{P(\overline{A} \cap B)}{P(\overline{A})} = \frac{0{,}40}{0{,}45} = \frac{8}{9} = 88{,}9\ \%$

Aufgabe 5

a)

$P(A) = \frac{5}{36}$ $(1 + 5 = 5 + 1 = 2 + 4 = 4 + 2 = 3 + 3 = 6)$;
$P(B) = \frac{6}{36} = \frac{1}{6}$ $(1 - 1 = 2 - 2 = 3 - 3 = 4 - 4 = 5 - 5 = 6 - 6 = 0)$;
$P(A \cap B)$ = P(Augensumme ist 6 **und** die Differenz ist Null) = P(beide Würfel zeigen eine 3)
$= \frac{1}{36} \neq P(A) \cdot P(B) = \frac{5}{216}$. Daher sind beide Ereignisse stochastisch abhängig.

b)

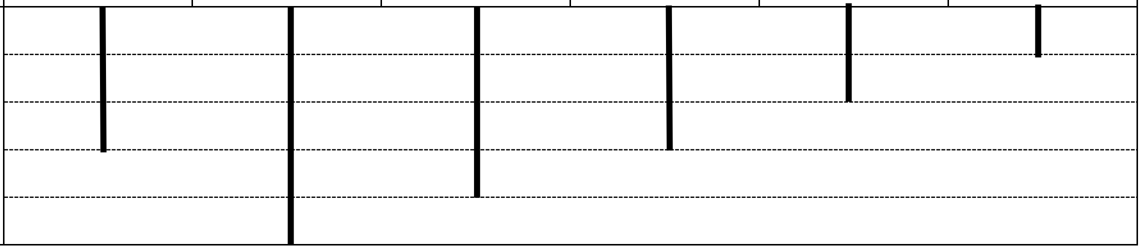
$P(A) = \frac{1}{6}$; $P(B) = \frac{36-8}{36} = \frac{7}{9}$ $(1 + 1, 1 + 2 = 2 + 1, 2 + 2, 2 + 3 = 3 + 2, 1 + 3 = 3 + 1$ sind kleiner als 6)
$P(A \cap B) = P((3/1)$ und $(3/2)) = \frac{2}{36} = \frac{1}{18} \neq P(A) \cdot P(B) = \frac{7}{63}$. A und B sind stochastisch abhängig.

c)

$P(A) = \frac{1}{3}$; $P(B) = \frac{1}{2}$; $P(A \cap B) = \frac{6}{36} = \frac{1}{6} = P(A) \cdot P(B)$. A und B sind stochastisch unabhängig.

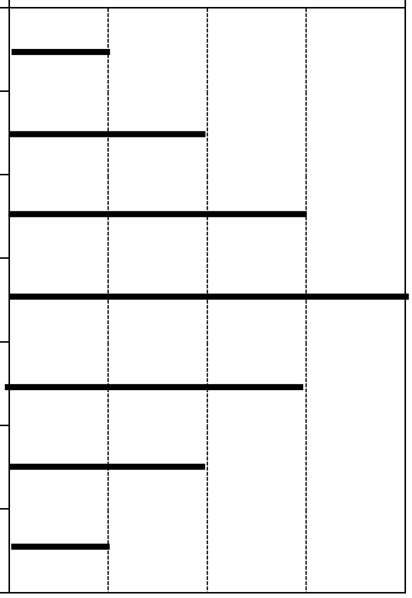
2 Wahrscheinlichkeitsverteilung und Histogramme

Aufgabe 1

x_i	0	1	2	3	4	5
$X = x_i$	(1/1) (2/2) (3/3) (4/4) (5/5) (6/6)	(1/2), (2/1) (2/3), (3/2) (3/4), (4/3) (4/5), (5/4) (5/6), (6/5)	(1/3), (3/1) (2/4), (4/2) (3/5), (5/3) (4/6), (6/4)	(1/4), (4/1) (2/5), (5/2) (3/6), (6/3)	(1/5), (5/1) (2/6), (6/2)	(1/6), (6/1)
$P(X = x_i)$	$\frac{6}{36} = \frac{1}{6} = \frac{3}{18}$	$\frac{10}{36} = \frac{5}{18}$	$\frac{8}{36} = \frac{2}{9} = \frac{4}{18}$	$\frac{6}{36} = \frac{1}{6} = \frac{3}{18}$	$\frac{4}{36} = \frac{1}{9} = \frac{2}{18}$	$\frac{2}{36} = \frac{1}{18}$
Histogramm $\frac{1}{18}$						

Der Erwartungswert beträgt $\mu = E(X) = 0 \cdot \frac{1}{6} + 1 \cdot \frac{5}{18} + 2 \cdot \frac{2}{9} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{9} + 5 \cdot \frac{1}{18} = \frac{35}{18} \approx 1{,}94$

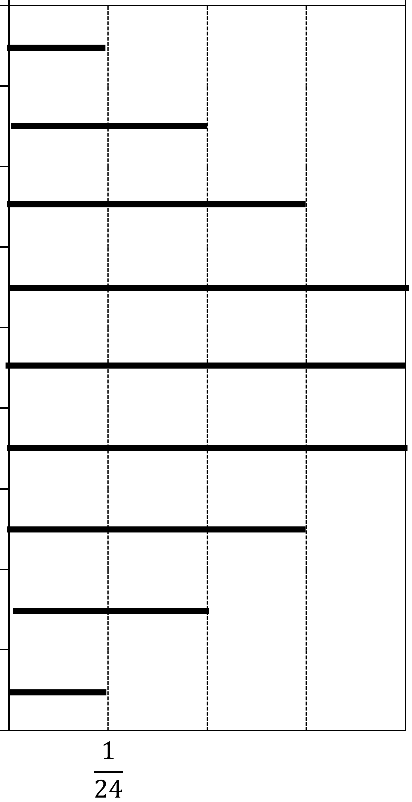
2a)

x_i	$X = x_i$: zugehörige Ergebnisse	$P(X = x_i)$	Histogramm
2	(1/1)	$\frac{1}{16}$	
3	(1/2), (2/1)	$\frac{2}{16} = \frac{1}{8}$	
4	(1/3), (3/1), (2/2)	$\frac{3}{16}$	
5	(1/4), (4/1), (2/3), (3/2)	$\frac{4}{16} = \frac{1}{4}$	
6	(2/4), (4/2), (3/3)	$\frac{3}{16}$	
7	(3/4), (4/3)	$\frac{2}{16} = \frac{1}{8}$	
8	(4/4)	$\frac{1}{16}$	

$\frac{1}{16}$

2b)

x_i	$X = x_i$: zugehörige Ergebnisse	$P(X = x_i)$	Histogramm
2	(1/1)	$\frac{1}{24}$	
3	(1/2), (2/1)	$\frac{2}{24} = \frac{1}{12}$	
4	(1/3), (3/1), (2/2)	$\frac{3}{24} = \frac{1}{8}$	
5	(1/4), (4/1), (2/3), (3/2)	$\frac{4}{24} = \frac{1}{6}$	
6	(1/5), (2/4), (4/2), (3/3)	$\frac{4}{24} = \frac{1}{6}$	
7	(1/6), (2/5), (3/4), (4/3)	$\frac{4}{24} = \frac{1}{6}$	
8	(2/6), (3/5), (4/4)	$\frac{3}{24} = \frac{1}{8}$	
9	(3/6), (4/5)	$\frac{2}{24} = \frac{1}{12}$	
10	(4/6)	$\frac{1}{24}$	



2c)

x_i	Oktaeder und Tetraeder		Zwei Hexaeder	
	$X = x_i$	$P(X = x_i)$	$X = x_i$	$P(X = x_i)$
2	(1/1)	$\frac{1}{32}$	(1/1)	$\frac{1}{36}$
3	(2/1), (1/2)	$\frac{1}{16}$	(2/1), (1/2)	$\frac{1}{18}$
4	(3/1), (2/2), (1/3)	$\frac{3}{32}$	(3/1), (2/2), (1/3)	$\frac{1}{12}$
5	(4/1), (3/2), (2/3), (1/4)	$\frac{1}{8}$	(4/1), (3/2), (2/3), (1/4)	$\frac{1}{9}$
6	(5/1), (4/2), (3/3), (2/4)	$\frac{1}{8}$	(5/1), (4/2), (3/3), (2/4), (1/5)	$\frac{5}{36}$
7	(6/1), (5/2), (4/3), (3/4)	$\frac{1}{8}$	(6/1), (5/2), (4/3), (3/4), (2/5), (1/6)	$\frac{1}{6}$
8	(7/1), (6/2), (5/3), (4/4)	$\frac{1}{8}$	(6/2), (5/3), (4/4), (3/5), (2/6)	$\frac{5}{36}$
9	(8/1), (7/2), (6/3), (5/4)	$\frac{1}{8}$	(6/3), (5/4), (4/5), (3/6)	$\frac{1}{9}$
10	(8/2), (7/3), (6/4)	$\frac{3}{32}$	(6/4), (5/5), (4/6)	$\frac{1}{12}$
11	(8/3), (7/4)	$\frac{1}{16}$	(6/5), (5/6)	$\frac{1}{18}$
12	(8/4)	$\frac{1}{32}$	(6/6)	$\frac{1}{36}$

Eine Häufigkeitsverteilung beim zweifachen Hexaeder-Wurf nimmt zur Mitte hin gleichmäßig zu und von dort gleichmäßig ab. Beim Wurf mit einem Oktaeder und Hexaeder gibt es ein breites mittleres Plateau.

2d)

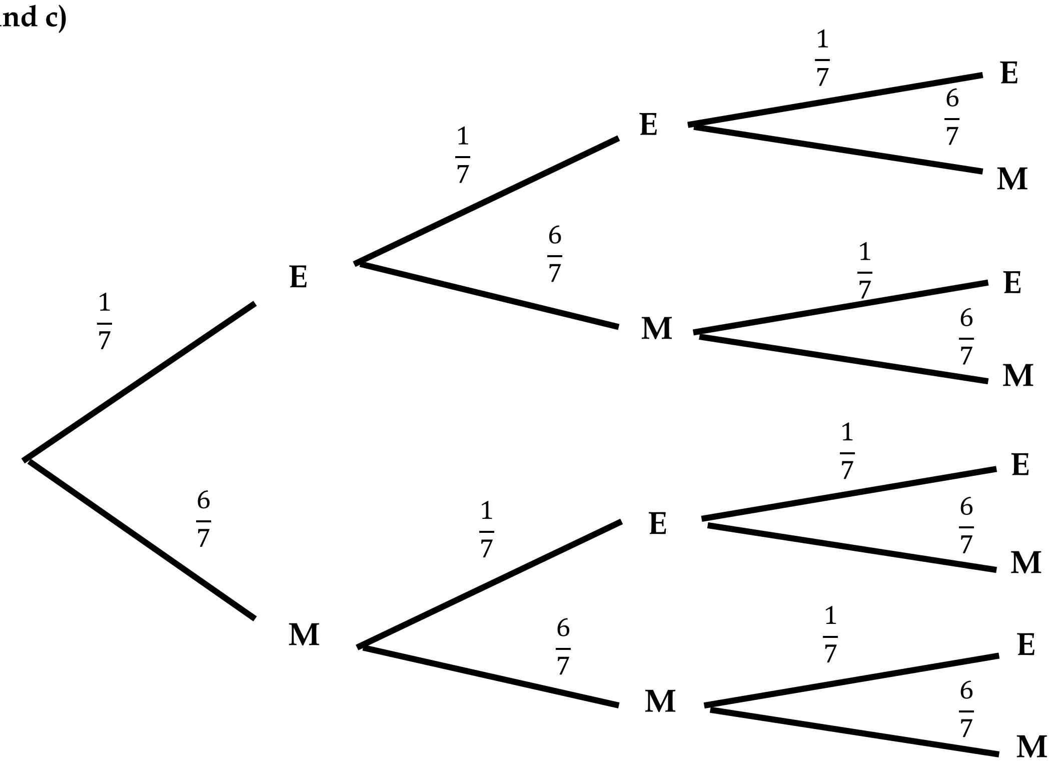
Versuch	Erwartungswert μ
2-facher Wurf eines Tetraeders	5
Wurf eines Tetraeders und Hexaeders	6
Wurf eines Oktaeders und Tetraeders	7
Wurf zweier Hexaeder	7

Die Erwartungswerte lassen sich mit der Wahrscheinlichkeitsverteilung berechnen oder mit folgender Überlegung: Beim Tetraeder erhält man beim Werfen eine Würfels im Durchschnitt langfristig (1 +2 + 3 + 4) : 4 = 2,5, beim Hexaeder (1 + 2+ …_+ 6) : 6 = 3,5 und beim Werfen des Oktaeders (1 + 2 + …+ 8) : 8 = 4,5. Durch Addition der beiden Prognosewerte erhält man die Erwartungswerte der obigen Tabelle.

3 Binomialverteilung

Aufgabe 1

a) und c)



b) Anzahl n der Schokoüberraschungskugeln = 2

Erfolgsanzahl k	X = k: mögliche Ergebnisse mit k Treffern	Wahrscheinlichkeit P(X = k)
0	(M/M)	$\left(\frac{6}{7}\right)^2 = 1 \cdot \left(\frac{1}{7}\right)^0 \cdot \left(\frac{6}{7}\right)^2 \approx 73{,}47\ \%$
1	(E/M), (M/E)	$2 \cdot \frac{1}{7} \cdot \frac{6}{7} = 2 \cdot \left(\frac{1}{7}\right)^1 \cdot \left(\frac{6}{7}\right)^1 \approx 24{,}48\ \%$
2	(E/E)	$\left(\frac{1}{7}\right)^2 = 1 \cdot \left(\frac{1}{7}\right)^2 \cdot \left(\frac{6}{7}\right)^0 \approx 2{,}04\ \%$

c) n = 3

K	X = k	P(X = k)
0	(M/M/M)	$\left(\frac{6}{7}\right)^3 = 1 \cdot \left(\frac{1}{7}\right)^0 \cdot \left(\frac{6}{7}\right)^3 \approx 62{,}67\ \%$
1	(E/M/M), (M/E/M), (M/M/E)	$3 \cdot \frac{1}{7} \cdot \frac{6}{7} \cdot \frac{6}{7} = 3 \cdot \left(\frac{1}{7}\right)^1 \cdot \left(\frac{6}{7}\right)^2 \approx 31{,}48\ \%$
2	(E/E/M), (E/M/E), (M/E/E)	$3 \cdot \frac{1}{7} \cdot \frac{1}{7} \cdot \frac{6}{7} = 3 \cdot \left(\frac{1}{7}\right)^2 \cdot \left(\frac{6}{7}\right)^1 \approx 5{,}24\ \%$
3	(E/E/E)	$\left(\frac{1}{7}\right)^3 = 1 \cdot \left(\frac{1}{7}\right)^3 \cdot \ldots \cdot \left(\frac{6}{7}\right)^0 \approx 0{,}29\ \%$

d)

K	**X = k**	**P(X = k)**
0	(M/M/M/M)	$\left(\frac{6}{7}\right)^4 = 1 \cdot \left(\frac{1}{7}\right)^0 \cdot \left(\frac{6}{7}\right)^4 \approx 53{,}97\ \%$
1	(E/M/M/M), (M/E/M/M), (M/M/E/M), (M/M/M/E)	$4 \cdot \frac{1}{7} \cdot \frac{6}{7} \cdot \frac{6}{7} \cdot \frac{6}{7} = 4 \cdot \left(\frac{1}{7}\right)^1 \cdot \left(\frac{6}{7}\right)^3 \approx 35{,}98\ \%$
2	(E/E/M/M), (E/M/E/M), (E/M/M/E) (M/E/E/M), (M/E/M/E), (M/M/E/E)	$6 \cdot \frac{1}{7} \cdot \frac{1}{7} \cdot \frac{6}{7} \cdot \frac{6}{7} = 3 \cdot \left(\frac{1}{7}\right)^2 \cdot \left(\frac{6}{7}\right)^2 \approx 8{,}99\ \%$
3	(M/E/E/E), (E/M/E/E), (E/E/M/E), (E/E/E/M)	$4 \cdot \frac{1}{7} \cdot \frac{1}{7} \cdot \frac{1}{7} \cdot \frac{6}{7} = 4 \cdot \left(\frac{1}{7}\right)^3 \cdot \left(\frac{6}{7}\right)^1 \approx 1\ \%$
4	(E/E/E/E)	$\left(\frac{1}{7}\right)^4 = 1 \cdot \left(\frac{1}{7}\right)^4 \cdot \left(\frac{6}{7}\right)^0 \approx 0{,}04\ \%$

e)

$$P(X = k) = \text{Anzahl der Pfade mit k Erfolgen} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{n-k}$$

f)

N	**Erwartungswert μ**
1	$0 \cdot \frac{6}{7} + 1 \cdot \frac{1}{7} = \frac{1}{7}$
2	$\frac{2}{7}$
3	$\frac{3}{7}$
4	$\frac{4}{7}$
N	$\frac{n}{7}$

Für eine Schokokugel ist der Erwartungswert offenbar $\frac{1}{7}$. Bei sieben Kugeln wird man durchschnittlich 1 Erfolg erwarten können. Daher ist der Erwartungswert für n Kugeln n-mal so groß wie für 1 Kugel. Es gilt der wichtige Satz (vgl. Arbeitsblatt 12):

Sei X die Zufallsgröße: „*Trefferanzahl k bei n Versuchen mit der Trefferwahrscheinlichkeit p*“, dann gilt für den Erwartungswert $\mu = n \cdot p$.

Aufgabe 2

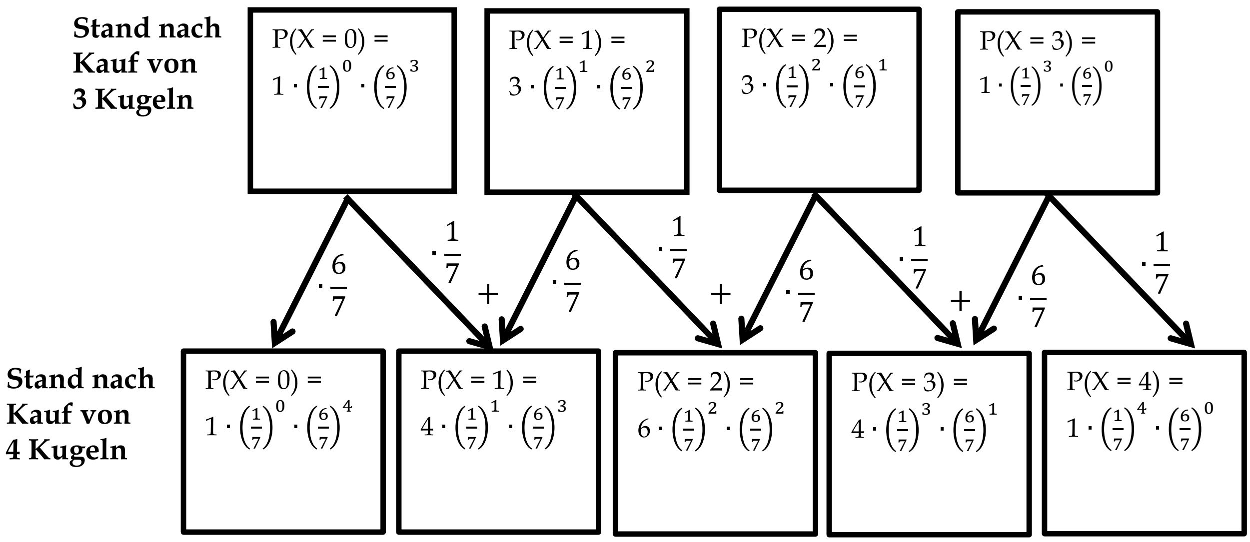
a)

- X: *„Anzahl der Treffer beim Elfmeterschießen bei n Spielen mit Trefferwahrscheinlichkeit p"*
- Y: *„Anzahl der HIV-Positiv-Erkrankungen in Deutschland bei einer Erkrankungswahrscheinlichkeit p"*
- Z: *„Anzahl der Erfolge beim Lotto 6 aus 49"*

b)

- Z: *„Anzahl der Erfolge beim Lotto 6 aus 49"*
- A: *„Anzahl der Treffer von der Freiwurflinie vor und nach einer Wurftrainingsphase"*
- B: *„Anzahl der Verspätungen vor und nach einem Fernsehbericht über die Pünktlichkeit der Bahn."*

c)



d)

Die Wahrscheinlichkeit der unteren Zeile setzt sich zusammen aus dem $\frac{1}{7}$- und oder $\frac{6}{7}$-fachen der oberen Wahrscheinlichkeiten. Treffen zwei Pfeile auf einen Kasten, werden dort die entsprechenden um den Faktor $\frac{1}{7}$ bzw. $\frac{6}{7}$ multiplizierten Ausgangswahrscheinlichkeiten addiert.

e)

P(X = 0) = $1 \cdot \left(\frac{1}{7}\right)^0 \cdot \left(\frac{6}{7}\right)^5$	P(X = 1) = $5 \cdot \left(\frac{1}{7}\right)^1 \cdot \left(\frac{6}{7}\right)^4$	P(X = 2) = $10 \cdot \left(\frac{1}{7}\right)^2 \cdot \left(\frac{6}{7}\right)^3$	P(X = 3) = $10 \cdot \left(\frac{1}{7}\right)^3 \cdot \left(\frac{6}{7}\right)^2$	P(X = 4) = $5 \cdot \left(\frac{1}{7}\right)^4 \cdot \left(\frac{6}{7}\right)^1$	P(X = 5) = $1 \cdot \left(\frac{1}{7}\right)^5 \cdot \left(\frac{6}{7}\right)^0$

f)

Die Summe der Exponenten muss stets 5 ergeben. Die Summe der Pfade ergibt $2^5 = 32$.

g) $\binom{10}{5} = 252$; $\binom{6}{4} = 15$; $\binom{10}{10} = 1$; $\binom{10}{1} = 10$; $\binom{100}{99} = 100$; $\frac{100!}{99!} = 100$

h)

Will man 3 Frauen und 2 Männer ohne Beachtung der Individualität der Personen auf 5 Stühle verteilen, überlegt man sich zuerst, wie viele Möglichkeiten es unter Beachtung der Individualität es gibt 5 Personen auf 5 Stühle zu verteilen. Dies sind 5! = 5 · 4 · 3 · 2 · 1 = 120 Möglichkeiten. Achtet man nicht mehr auf die Individualität der Frauen, können jeweils 3! = 3 · 2 · 1 = 6 Möglichkeiten für die Frauen zu einer Möglichkeit zusammengefasst werden. Man erhält 120 : 6 = 20 Möglichkeiten. Unter diesen Möglichkeiten können bei Nichtbeachtung der Individualität der Männer 2! = 2 · 1 = 2 erneut zusammengefasst werden. Man erhält insgesamt 20 : 2 = 10 Möglichkeiten. Es gilt:

$$\binom{5}{3} = \frac{\text{Möglichkeitsanzahl,5 Personen auf 5 Stühle zu verteilen}}{(\text{Möglichkeitsanzahl,3 Frauen auf 3 Stühle zu verteilen})\cdot(\text{Möglichkeitsanzahl,2 Männer auf zwei Stühle zu verteilen})}$$
$$= \frac{5!}{3!\cdot(5-3)!} = \frac{120}{6\cdot 2} = 10$$

Alternative Überlegungen: Unter Beachtung der Individualität gibt es 5 · 4 · 3 = 60 Möglichkeiten, 3 Frauen auf 5 Stühle zu verteilen. Achtet man nun nicht auf die Individualität, können 3 · 2 · 1 = 6 Möglichkeiten zu einer Möglichkeit zusammengefasst werden. So ergeben sich 60 : 6 = 10 Möglichkeiten. Es gilt:

$$\binom{5}{3} = \frac{\text{Möglichkeitsanzahl,3 Frauen unter Beachtung der Individualität auf 5 Stühle zu verteilen}}{\text{Möglichkeitsanzahl,3 Frauen auf 3 Stühle zu verteilen}} = \frac{5\cdot 4\cdot 3}{3!} = \frac{60}{6} = 10$$

i)

Die Anzahl der Pfade $\binom{n}{k}$, die zu Trefferzahl k bei n Versuchen führt, lässt sich ablesen als Eintrag im Pascalschen Dreieck, indem man in der n-ten Zeile den k-ten Eintrag von links sucht. Allgemein gilt das Schema wie bereits oben in den Aufgabenteilen c) und d) exemplarisch beschrieben:

P (X = k – 1; n – 1 Versuche) · p + (X = k; n – 1 Versuche) · (1 – p) = P(X = k; n Versuche)

Dies muss rechnerisch nachgewiesen werden:

$P_{n-1}(X = k - 1) \cdot p + P_{n-1}(X = k - 1) \cdot (1 - p)$

$= \binom{n-1}{k-1} \cdot p^{k-1} \cdot (1-p)^{n-k} \cdot p + \binom{n-1}{k} \cdot p^k \cdot (1-p)^{n-1-k} \cdot (1-p)$

$= [\binom{n-1}{k-1} + \binom{n-1}{k}] \cdot p^k \cdot (1-p)^{n-k}$

$= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$

Denn es gilt:

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \frac{(n-1)!}{(k-1)!\cdot(n-k)!} + \frac{(n-1)!}{k!\cdot(n-1-k)!} = \frac{(n-1)!\cdot k}{k!\cdot(n-k)!} + \frac{(n-1)!\cdot(n-k)}{k!\cdot(n-k)!}$$

$$= \frac{(n-1)!\cdot k + (n-1)!\cdot(n-k)}{k!\cdot(n-k)!} = \frac{(n-1)!\cdot k + (n-1)!\cdot n - (n-1)!\cdot k}{k!\cdot(n-k)!} = \frac{n!}{k!\cdot(n-1)!} = \binom{n}{k}$$

Aufgabe 3

Auslastungsmodell

a) In einem Drittel der Arbeitszeit wird eine Maschine benötigt.

b) Für einen Mitarbeiter handelt es sich um einen Bernoulli-Versuch mit der Maschinennutzungswahrscheinlichkeit $p = \frac{1}{3}$, da er entweder eine Maschine nutzt oder nicht. Da für jeden weiteren Mitarbeiter die entsprechende Maschinennutzungswahrscheinlichkeit unverändert bei $p = \frac{1}{3}$ bleibt, erhält man bei fünf Mitarbeitern eine Bernoulli-Kette der Länge 5.

c) $P(X = 0) = \binom{5}{0} \cdot \left(\frac{1}{3}\right)^0 \cdot \left(\frac{2}{3}\right)^5 \approx 13{,}17\ \%$ $P(X = 1) = \binom{5}{1} \cdot \left(\frac{1}{3}\right)^1 \cdot \left(\frac{2}{3}\right)^4 \approx 32{,}92\ \%$

d) $P(X = 2) = \binom{5}{2} \cdot \left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^3 \approx 32{,}92\ \%$ $P(X = 3) = \binom{5}{3} \cdot \left(\frac{1}{3}\right)^3 \cdot \left(\frac{2}{3}\right)^2 \approx 16{,}46\ \%$

$P(X = 4) = \binom{5}{4} \cdot \left(\frac{1}{3}\right)^4 \cdot \left(\frac{2}{3}\right)^1 \approx 4{,}12\ \%$ $P(X = 5) = \binom{5}{4} \cdot \left(\frac{1}{3}\right)^4 \cdot \left(\frac{2}{3}\right)^1 \approx 0{,}41\ \%$

e) Berechne die kumulierte Wahrscheinlichkeit $P(X > k)$ mit $k = 2, 3, 4$:

$P(X > 2) = 16{,}46\ \% + 4{,}12\ \% + 0{,}41\ \% = 20{,}99\ \%$ (Wenn drei Arbeiter zu einem bestimmten Zeitpunkt an einer Maschine sind, reichen zwei Maschinen nicht mehr aus. Es werden mindestens drei Maschinen benötigt.)

$P(X > 3) = 4{,}12\ \% + 0{,}41\ \% = 4{,}53\ \%$ (Wenn vier Arbeiter zu einem bestimmten Zeitpunkt an einer Maschine sind, reichen drei Maschinen nicht mehr aus. Es werden mindestens vier Maschinen benötigt.)

$P(X > 4) = 0{,}41\ \%$

Bei zwei Maschinen liegt die Wartezeit bei knapp 21 %, bei drei Maschinen sinkt sie unter 5 % und bei vier Maschinen unter 1 %.

f) Die Anschaffung einer weiteren Maschine macht Sinn, da beim Kauf einer dritten Maschine die Wahrscheinlichkeit einer Wartezeit um mehr als 16 % sinkt, während die Wahrscheinlichkeit bei der Anschaffung einer vierten Maschine nur noch um weitere knapp 5 % verringert wird.

Das Kugel-Fächer-Modell

a) Es handelt sich um einen Bernoulli-Versuch der Länge 150 mit der Geburtstagswahrscheinlichkeit von $p = \frac{1}{365}$. Die Zufallsgröße X lautet: X: *„Anzahl der Geburtstagskinder an einem beliebigen Tag"*. Die gesuchten Wahrscheinlichkeiten lassen sich mithilfe der Bernoulli-Formel berechnen:
$P(X = 0) = \binom{150}{0} \cdot \left(\frac{1}{365}\right)^0 \cdot \left(\frac{364}{365}\right)^{150} \approx 66{,}26\ \%$ $\quad P(X = 1) = \binom{150}{1} \cdot \left(\frac{1}{365}\right)^1 \cdot \left(\frac{364}{365}\right)^{149} \approx 27{,}31\ \%$
$P(X = 2) = \binom{150}{2} \cdot \left(\frac{1}{365}\right)^2 \cdot \left(\frac{364}{365}\right)^{148} \approx 5{,}59\ \%$

b) Um herauszufinden, wie oft im Jahr unter 150 Personen an einem beliebigen Tag keine Person Geburtstag hat, muss die Wahrscheinlichkeit, dass an einem beliebigen Tag niemand der 150 Personen Geburtstag hat, mit der Anzahl der Tage eines Jahres (365) multipliziert werden. So erhält man die Anzahl der Tage ohne Geburtstagskind. An $P(X = 0) \cdot 365 \approx 0{,}6626 \cdot 365 \approx 242$ Tagen eines Jahres hat im Schnitt also *keine* Person Geburtstag. An $P(X = 1) \cdot 365 \approx 100$ Tagen hat im Schnitt *genau eine* Person Geburtstag. Daher hat an 365 – 242 – 100 = 23 Tagen im Schnitt eine Person mit *mindestens einer weiteren* Person Geburtstag.

c) Wie in Aufgabenteil a) handelt es sich um eine 150-stufigen Bernoulli-Kette mit der Trefferwahrscheinlichkeit von $p = \frac{1}{365}$. Gesucht ist die durchschnittliche Anzahl von Zahlen unter den Zahlen 1, 2, 3, …, 365, die nicht unter den 150 gezogenen ist. Dafür multipliziert man wie in Aufgabenteil b) die Bernoulli-Wahrscheinlichkeit $P(X = 0)$ (Wahrscheinlichkeit, dass keine der 150 Zahlen gezogen wird) mit 365 und erhält 241,85. Von 365 Zahlen kommen im Schnitt also 242 Zahlen *nicht* vor. Multipliziert man $P(X = 1)$ (Wahrscheinlichkeit, dass eine Zahl genau einmal vorkommt) mit 365, ergibt sich 99,68. Hier kann man sagen: Von 365 Zahlen kommen im Schnitt 100 Zahlen *genau einmal* vor. Daher kommen 365 – 242 – 100 = 23 Zahlen *mehr als einmal* vor.

d) Die Anzahl der Kugeln ist n = 150 und entspricht der Anzahl der erzeugten Zufallszahlen. Die Anzahl der Fächer beträgt f = 365 und entspricht den Zahlen 1, 2, 3, …, 365. Ferner sei k = 0, 1, …, 150 die Häufigkeit, mit der eine Kugel in einem bestimmten Fach landet. Dieser Parameter entspricht der absoluten Häufigkeit, mit der eine Zahl gezogen wird. Die Zufallsgröße lautet daher: *„Häufigkeit, mit der eine Kugel in einem Fach landet"*. Sie entspricht der Zufallsgröße: *„Häufigkeit, mit der eine Zahl gezogen wird"*. Aus der Sicht eines Faches beträgt die Wahrscheinlichkeit, dass eine der 150 Kugeln (Zufallszahlen) in ihm abgelegt wird, $p = \frac{1}{365}$. Bei n = 150 Kugeln (Anzahl der Zufallszahlen) erhält man daher eine 150-stufige Bernoulli-Kette mit der Trefferwahrscheinlich $p = \frac{1}{365}$. Es ergibt sich allgemein für das Kugel-Fächer-Modell mit den Parametern Kugelanzahl (Anzahl der Zufallszahlen) n, Fächerbelegungszahl k (Häufigkeit, mit der eine Zahl gezogen wird) und Anzahl der Fächer f (Zahlen 1, 2, 3, …) mithilfe der Bernoulli-Formel: $P(X = k) = \binom{n}{k} \cdot \left(\frac{1}{f}\right)^k \cdot \left(1 - \frac{1}{f}\right)^{n-k}$.

e) Die Anzahl der Rosinen ist n = 100. Die Anzahl der Brötchen beträgt f = 6. Die Wahrscheinlichkeit, dass eine Rosine in einem bestimmten Brötchen landet beträgt daher $p = \frac{1}{6}$. Gesucht ist die Wahrscheinlichkeit, dass bei n = 100 Rosinen ein Brötchen ohne Rosinen ist. Diese Wahrscheinlichkeit entspricht $P(X = 0) = \binom{100}{0} \cdot \left(\frac{1}{6}\right)^0 \cdot \left(\frac{5}{6}\right)^{100} \approx 0{,}000000012 = 0{,}0000012\ \%$. Für die weiteren zu untersuchenden Fälle gilt: $P(X = 5) \approx 0{,}00029 = 0{,}029\ \%$; $P(X = 20) \approx 10{,}65\ \%$; $P(X = 20) \approx 6{,}8\ \%$; $P(X \geq 16) \approx 61{,}23\ \%$; $P(10 \leq X \leq 20) \approx 82{,}68\ \%$.

Das Geburtstagsparadoxon

a) Bei beiden Fragestellungen handelt es sich um ein Bernoulli-Experiment, da es in beiden Versuchen nur zwei mögliche Ausgänge des Versuchs gibt (Geburtstagsproblem: Person hat mit einer anderen Person Geburtstag oder nicht / Würfelproblem: Würfel zeigt eine Augenzahl eines zweiten Würfels oder nicht)

Allerdings entsteht keine Bernoulli-Kette, da sich die Trefferwahrscheinlichkeit mit höherer Versuchszahl verändert. Z. B. wäre die Trefferwahrscheinlichkeit beim Geburtstagsproblem bei 367 Personen (unter Berücksichtigung des Schaltjahres) für die letzte Person 1, während die gleiche Wahrscheinlichkeit bei 2 Personen $\frac{1}{365}$ beträgt. Ebenso wird beim 7-stufigen Hexaeder-Wurf der siebte Wurf eine Trefferwahrscheinlichkeit von 1 haben, während die Trefferwahrscheinlichkeit des zweiten Wurfes $\frac{1}{6}$ beträgt.

b) Statt die Wahrscheinlichkeiten für alle Ergebnisse mit *„mindestens doppelt auftretender Augenzahl"* zu betrachten, bestimmt man die Wahrscheinlich des Gegenereignisses, das nur aus einem Ergebnis besteht, nämlich dem Ereignis, dass *„alle Augenzahlen ungleich"* sind. Die gesuchte Wahrscheinlichkeit ergibt sich durch Subtraktion der Wahrscheinlichkeit des Gegenereignisses von 100 % = 1.

$$\boxed{\text{eine Augenzahl}} \xrightarrow{\frac{5}{6}} \boxed{\text{zwei verschiedene Augenzahlen}} \xrightarrow{\frac{4}{6}} \boxed{\text{drei verschiedene Augenzahlen}} \xrightarrow{\frac{3}{6}} \boxed{\text{vier verschiedene Augenzahlen}}$$

P (*alle vier Würfel haben eine unterschiedliche Augenzahl*) = $= \frac{5}{6} \cdot \frac{4}{6} \cdot \frac{3}{6} = \frac{5 \cdot 4 \cdot 3}{6^3} = \frac{5}{18} \approx 27{,}8\ \%$. Damit beträgt die Wahrscheinlichkeit für das Ereignis *„mindestens eine doppelt auftretende Augenzahl"* $1 - \frac{5}{18} = \frac{13}{18} \approx 72{,}2\ \%$.

c) Das hängt davon ab, wie die Auszahlung a im Verhältnis zum Einsatz für den Fall wäre, dass mindestens eine Augenzahl doppelt vorkommt. Für die folgende Wahrscheinlichkeitsverteilung mit der Zufallsgröße X: *„Auszahlung pro Spiel in €"* ergibt sich die Wahrscheinlichkeitsverteilung für X:

K	*Alle 4 Augenzahlen ungleich*	*mindestens eine doppelt auftretende Augenzahl*
Auszahlung	0	A
P(X = k)	$\frac{5}{18}$	$\frac{13}{18}$

Im Falle des Einsatzes von 1 € müsste der Erwartungswert $\mu = E(X) = 0 \cdot \frac{5}{18} + a \cdot \frac{13}{18} = 1$ betragen, was bedeutet, dass die Auszahlung $a = \frac{18}{13} \approx 1{,}38$ €. Bei einer geringeren Auszahlung würde man nicht auf das Ereignis *„mindestens eine doppelt auftretende Augenzahl"* wetten.

d) P(*mindestens 2 Personen haben am gleichen Tag Geburtstag*) = 1 – P (*alle Personen haben an verschiedenen Tagen Geburtstag*) $= 1 - \frac{364 \cdot 363 \cdot 362 \cdot \ldots \cdot 343}{365^{22}} = 1 - \frac{\binom{364}{22} \cdot 22!}{365^{22}} \approx 0{,}5073 = 50{,}73\ \%$.

e) (e.1) Gesucht ist die Wahrscheinlichkeit, dass *eine Person mindestens zwei Preise* bekommt. Berechne zunächst die Wahrscheinlichkeit für das Ereignis, dass *jede Person höchstens einen Preis* erhält. Sie beträgt $\frac{24\cdot 23\cdot\ldots\cdot 16}{25^9} = \frac{\binom{24}{9}\cdot 9!}{25^9} \approx 0{,}1244 = 12{,}44\ \%$. Daher beträgt die Wahrscheinlichkeit dafür, dass *eine Person mindestens zwei Preise* erhält etwa 1 – 12,44 % = 87, 46 %.

(e.2) Auch hier betrachtet man zunächst das Gegenereignis „*zehn unterschiedliche Zahlen in Folge*“. Die Wahrscheinlichkeit dafür beträgt $\frac{36\cdot 35\cdot\ldots\cdot 26}{37^9} = \frac{\binom{36}{9}\cdot 9!}{37^9} \approx 0{,}2629 = 26{,}29\ \%$. Daher beträgt die Wahrscheinlichkeit für das Gegenereignis „mindestens eine doppelte Zahl“ 73,71 %.

(e.3) P (*fünf unterschiedliche Motive*) = $\frac{39\cdot 38\cdot 37\cdot 36}{40^4} \approx 77{,}11\ \%$. Daher P (*mindestens ein doppeltes Motiv*) = 1 – 77,11 % = 22,89 %.

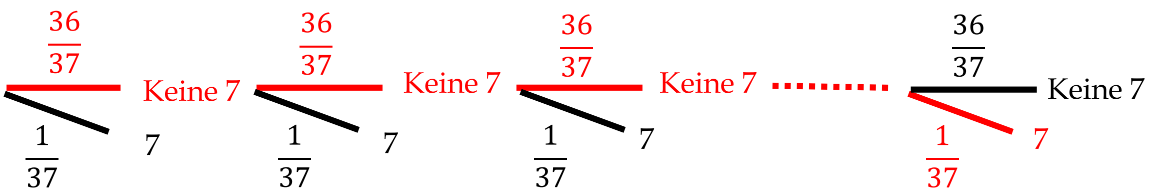
(e.4) P (*mindestens zwei Druckfehler auf einer Seite*) = 1 – P (*höchstens ein Druckfehler pro Seite*) = $1 - \frac{99\cdot 98\cdot 97\cdot\ldots\cdot 91}{100^9} = 1 - \frac{\binom{99}{9}\cdot 9!}{100^9} \approx 1 - 0{,}6282 = 37{,}18\ \%$.

(e.5) P (*mindestens zwei Rosinen pro Brötchen*) = 1 – P (*höchstens 1 Rosine pro Brötchen*) = $1 - \frac{49\cdot 48\cdot 47\cdot\ldots\cdot 31}{50^{19}} = 1 - \frac{\binom{49}{19}\cdot 19!}{50^{19}} \approx 1 - 0{,}0102 = 98{,}80\ \%$.

(e.6) P (*Jede Ente wird von höchstens einem Jäger abgeschossen*) = $\frac{19\cdot 18\cdot 17\cdot 16}{20^4} = 58{,}14\ \%$. Das Gegenereignis wird hier nicht benötigt. Man geht davon aus, dass die Jäger auf jeden Fall treffen. Trifft der erste Jäger eine Ente, bleiben für den zweiten nur noch 19 Enten übrig, wenn er nicht die gleiche Ente treffen will. Für den dritten Jäger bleiben 18, für den vierten Jäger 17 und für den fünften Jäger 16 Enten übrig, wenn jeweils eine andere Ente getroffen werden soll. Da alle gleichzeitig und unabhängig voneinander schießen haben alle fünf Jäger 20 Enten zur Auswahl.

Warten auf Erfolg

a) Aus dem Baumdiagramm kann man ablesen: Für die Zufallsgröße X: *„Anzahl k der notwendigen Runden, bis die Kugel auf Feld ‚7' liegen bleibt"*, gilt: $P(X = k) = \frac{1}{37} \cdot \left(\frac{36}{37}\right)^{k-1}$.



b) Dass die Kugel keinmal in k Runden auf Feld „7" liegen bleibt, bedeutet: lauter Misserfolge in k Runden. Die Wahrscheinlichkeit hierfür beträgt: $P(X > k) = \left(\frac{36}{37}\right)^{k}$ (Die Anzahl der notwendigen Runden ist größer als k)

c) Lösung über das Gegenereignis: Sei X die Anzahl der Erfolge in n Runden und die Trefferwahrscheinlichkeit $p = \left(\frac{1}{37}\right)$ und der Nichttrefferwahrscheinlichkeit $q = \left(\frac{36}{37}\right)$ Wenn $P(X \geq 1) \geq 0{,}95$ sein soll, ist dies gleichbedeutend mit der Ungleichung $P(X = 0) \leq 0{,}05$. Die Wahrscheinlichkeit, keinen Treffer in n Versuchen zu erhalten, lässt sich berechnen durch $\left(\frac{36}{37}\right)^{n}$ berechnen. Daher ergibt sich die Ungleichung $\left(\frac{36}{37}\right)^{n} \leq 0{,}05$. Durch Logarithmieren erhält man (Hinweis zum GTR: MATH → logab): $\log_{\frac{36}{37}}\left(\left(\frac{36}{37}\right)^{n}\right) \leq \log_{\frac{36}{37}}(0{,}05) \Leftrightarrow n \geq \log_{\frac{36}{37}}(0{,}05) \approx 109{,}34$ (Der Logarithmus zur Basis $\frac{36}{37} < 1$ ist eine streng monoton fallende Funktion. Daher dreht sich bei der Ungleichung nach dem Logarithmieren das Zeichen „≤" um.). Auch Ausprobieren führt zum Ziel:

n	50	100	125	**110**	**109**
$\left(\frac{36}{37}\right)^{n}$	0,2541	0,0646	0,0325	**0,0491**	**0,0505**

Wenn man also mindestens 110 Runden beim Roulette-Spiel durchführt, wird die Kugel mit einer Wahrscheinlichkeit von mindestens 95% mindestens einmal im Feld „7" liegen bleiben.

Lösung über die kumulierte Wahrscheinlichkeit: Sei X die Anzahl der Erfolge in n Runden. Die Wahrscheinlichkeit für *„mindestens einen Erfolg in n Runden"* entspricht der kumulierten Wahrscheinlichkeit $P(X \geq 1)$. Gesucht ist bei $k \geq 1$ die Zahl n, so dass $P(X \geq 1) \geq 0{,}95$. Mithilfe der Tabellenfunktion (Menu 7) kann durch Eingabe von $y = \text{BinomialCD}(1, x, x, \frac{1}{37})$ herausgefunden werden, dass x größer als 109 sein muss.

Aufgabe 4

(1) p = 0,3

$P(X = 14) = 0{,}1189$
$P(X \le 18) = 0{,}8594$
$P(X < 13) = P(X \le 12) = 0{,}2229$ (höchstens 12 Treffer)
$P(X \ge 17) = P(X = 17, 18, \ldots, 50) = P(X \ge 17) = P(17 \le X \le 50) = 0{,}3161$ (mindestens 17 Treffer)
$P(X > 21) = P(X = 22, 23, \ldots, 50) = P(X \ge 22) = P(12 \le X \le 50) = 0{,}0251$ (mindestens 22 Treffer)
$P(10 \le X \le 18) = 0{,}8192$ (mindestens 10 und höchstens 18 Treffer)

(2) p = 0,6

$P(X = 31) = 0{,}1109$ (genau 31)
$P(X \le 33) = 0{,}8439$ (höchstens 33)
$P(X < 26) = P(X = 0, 1, 2, \ldots, 25) = 0{,}0978$ (höchstens 25)
$P(X \ge 27) = P(27 \le X \le 50) = 0{,}8438$ (mindestens 27)
$P(X > 31) = P(X \ge 32) = P(32 \le X \le 50)= 0{,}3356$ (mindestens 32)
$P(23 \le X \le 33) = 0{,}8279$ (mindestens 23 und höchstens 33)

(3) P(X) durch P(Y) ausdrücken

$P(X = 31) = \binom{50}{31} 0{,}6^{31}\, 0{,}4^{19} = \frac{50!}{31! \cdot 19!} 0{,}6^{31}\, 0{,}4^{19} = \frac{50!}{19! \cdot 31!} 0{,}4^{19} 0{,}6^{31} = \binom{50}{19} 0{,}4^{19} 0{,}6^{31} = P(Y = 19)$
$P(X \le 33) = P(0, 1, \ldots, 33) = P(Y = 50, 49, \ldots, 17) = P(Y \ge 17)$
$P(X < 26) = P(X = 0, 1, 2, \ldots, 25) = P(Y = 50, 49, 48, \ldots, 25)$
$P(X \ge 27) = P(X = 27, 28, 29, \ldots 50) = P(Y = 23, 22, \ldots, 0) = P(Y \le 23)$
$P(X > 31) = P(X \ge 32) = P(X = 32, 33, \ldots, 50) = P(Y = 18, 17, 16, \ldots, 0) = P(Y \le 18)$
$P(23 \le X \le 33) = P(X = 23, 24, \ldots, 33) = P(Y = 27, 26, \ldots, 17) = P(17 \le Y \le 27)$

Aufgabe 5

Grundaufgabe 1
a) $P(X = 5) \approx 0{,}1789$
b) $P(X \geq 9) \approx 0{,}3759$ $[P(5 \leq X \leq 8) \approx 0{,}6178]$

Grundaufgabe 2
a) Erstelle mit dem GTR in MENU 1 über BinomialPD(10, 0.4) eine Liste für die singulären Wahrscheinlichkeiten $P(X = k)$ mit $k = 0, 1, \ldots, 10$. Den Tabellenwerten entnimmt man den Höchstwert für $k = 4$, denn $P(X = 3) \approx 0{,}2150 < P(X = 4) \approx 0{,}2508 < P(X = 5) \approx 0{,}2007$.
b) Ansatz: Für welche k ist $P(X < k) = P(X \leq k - 1) \geq 0{,}95$. Lösung in MENU 7 (Tabellenfunktion): Man bestimmt für y = BinomialCD(x, 200, 0.02) die erste natürliche Zahl x (denke an SET und den Startwert 0, z. B. den Endwert 10 und die Schrittweite 1), so dass y mindestens 0,95 beträgt ist. Dies ist die Zahl $x = k - 1 = 7$. Fazit: Bei $k = 8$ nicht-normgerechter Schrauben sollte die Lieferung zurückgewiesen werden. Alternativ in MENU 1 (Rechenmodus): Bei einem nicht zu hohen Stichprobenumfang n kann über die Funktion BinomialCD(200, 0.02) eine Liste bestimmt werden für die kumulierten Wahrscheinlichkeiten $P(X \leq k - 1)$ mit $k = 1, 2, \ldots, 201$. Man erhält für $k - 1 = 7$ zu ersten Mal einen Wahrscheinlichkeitswert, der größer als 0,95 ist. Alternativ gelangt man auch mit dem Befehl InvBinomialCD(0.95, 200, 0.02) = 7 zur ersten Zahl $k - 1$, so dass $P(X \leq k - 1) > 0{,}95$.

Grundaufgabe 3
a) $P(X = 5) = \binom{5}{5} \cdot p^5 \cdot (1 - p)^0 \geq 0{,}95 \Leftrightarrow p^5 \geq 0{,}95 \Leftrightarrow p \geq \sqrt[5]{0{,}95} \approx 0{,}9897 = 98{,}97\ \%$
(einfacher Sonderfall: Im Bernoulli-Term haben der Binomialkoeffizient und die zweite Potenz jeweils den Wert 1).
b) Ansatz: Für welche Trefferwahrscheinlichkeit p ist $P(X \geq 38) \geq 0.9$. Lösungsmöglichkeit über MENU 5 (Grafenmenu): Betrachte die Funktionen y = BinomialCD(38, 40, 40, x) und y = 0,9 für den x-Bereich (x für die Trefferwahrscheinlichkeit p) und y-Bereich (y für die Wahrscheinlichkeit P) zwischen 0 und 1 (über V-Window). Ermittle grafisch den Schnittpunkt der beiden Graphen (über G-Solv und IINTSECT). Man erhält S (0,9721/0,9). Daher ist für $p \geq 0{,}9721$ $P(X \geq 38) \geq 0.9$.

Grundaufgabe 4
a) Für welche Zahl n ist $P(X \geq 1) \geq 0.9$? Über MENU 7 (Tabellenfunktion): Erstelle ein Tabelle für y = BinomialCD(1, x, x, 0.05) mit x = 1, 2, 3, … (x ist der Stichprobenumfang n und kann über SET mit Start- und Endwert sowie Schrittweite 1 eingestellt werden). Für x = 45 wird zum ersten Mal die Wahrscheinlichkeit von 0,90 überschritten. Daher müssen mindestens 45 Personen kontrolliert werden, damit ein Kontrolleur mindestens 1 Schwarzfahrer mit einer Wahrscheinlichkeit von mindestens 90 % erwischt. Rechnerische Alternativlösung über das Gegenereignis von $X \geq 1$: $P(X \geq 1) = 1 - P(X = 0) \geq 0.9$ bedeutet, dass $P(X = 0) = \binom{n}{0} \cdot 0{,}05^0 \cdot 0{,}95^n \leq 0.1$. Also erhält man: $0{,}95^n \leq 0.1 \Leftrightarrow n \geq \log_{0{,}95} 0{,}1 \approx 44{,}89$. Alternativ kann man auch unter MENU A (Gleichungsmenu) mit SOLVER arbeiten. Man gibt dort die Gleichung $0{,}95^n = 0{,}1$ ein und wählt den Bereich, in dem die Lösung liegen kann (unterer Wert 40, oberer Wert 50).
b) $P(X \geq 1) \geq 0.9$ ist gleichbedeutend mit $P(X = 0) = 0{,}65^n \leq 0{,}1$. Wie unter a) ergibt sich $n \geq 6$.
c) Für welches n ist $P(X \geq 3) \geq 0.95$? MENU 7 (Tabellenfunktion): Erstelle eine Tabelle für y = BinomialCD(3, x, x, 0.80) mit x = 3, 4, 5 … (x ist der Stichprobenumfang n und kann über SET mit Start- und Endwert sowie Schrittweite 1 eingestellt werden). Die Wahrscheinlichkeit 0,95 wird für n = 6 überschritten. Alternativ kann in MENU 1 (Rechenmodus) durch händische Eingabe von BinomialCD(3, x, x, 0.80) mit x = 4, 5, 6, … überprüft werden, wann diese Wahrscheinlichkeit > 0,95 ist.

Aufgabe 6

a)

Es sind bei 60 Würfen beim ungezinkten Würfel $\frac{1}{6} \cdot 60 = 10$ Einser zu erwarten. Sollten nur 3 Einser kommen, könnte es sich um einen gezinkten Würfel handeln, da der Wert sehr weit vom Erwartungswert abweicht.

b)

Die Formel für den Erwartungswert lautet $\mu = E(X) = n \cdot p$. Bei einem Wurf kommt zu einem Sechstel die „1". Bei 6 Würfen kommt im Schnitt einmal ($\frac{1}{6} \cdot 6 = 1$) die „1". Bei 60 Würfen beträgt die durchschnittlich zu erwartene Einserzahl $\frac{1}{6} \cdot 60 = 10$. Allgemein beträgt die Prognose für den zu erwartene Mittelwert $\frac{1}{6} \cdot n$. Für eine beliebige Trefferwahrscheinlichkeit erhält man die obige Formel.

c)

Erste Beobachtungen, z. B. (vgl. Aufgabe d)): Maximum bei steigender Wahrscheinlichkeit weiter rechts, bei größerer Zahl n nimmt das Maximum ab, für p = 0,5 ist das Histogramm symmetrisch zum Erwartungswert, bei großem n werden die Histogramm zunehmend symmetrisch zu Erwartungswert.

Erwartungswerte

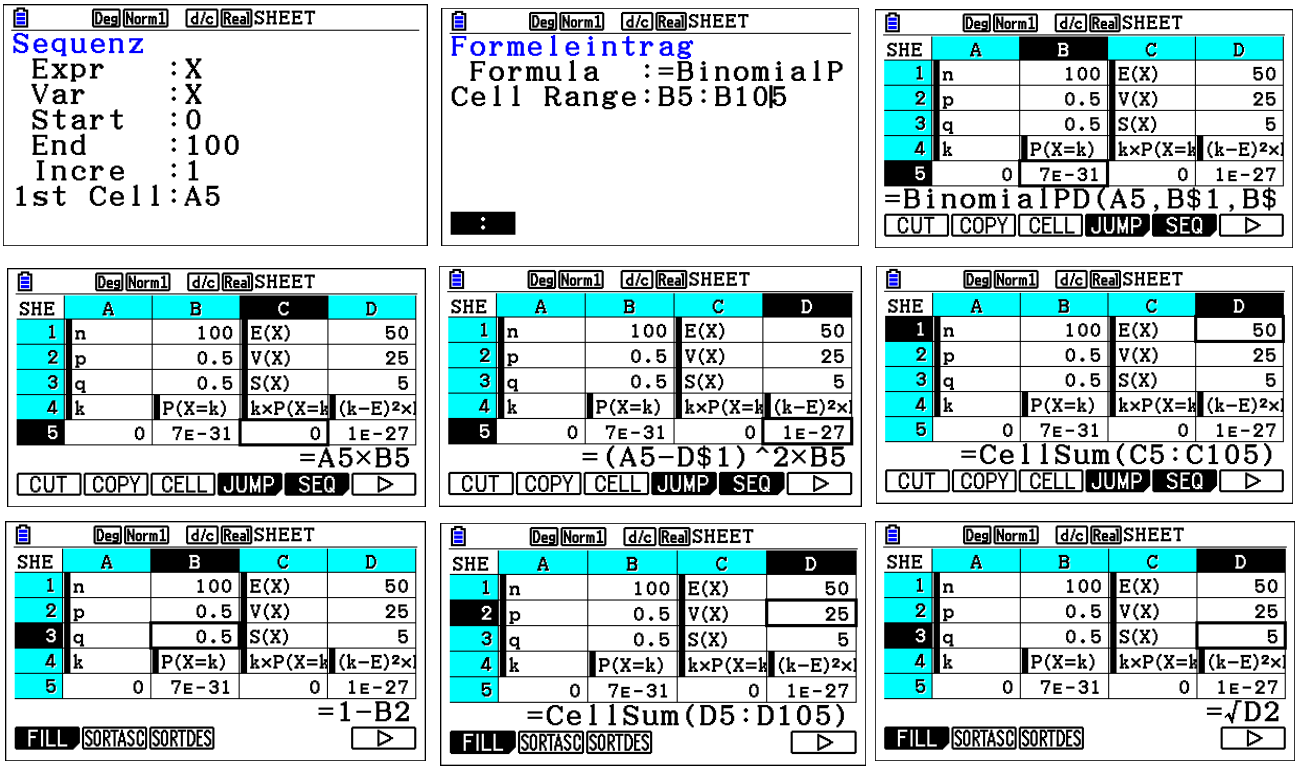
p	n = 4	n = 10	n = 20
0,1	0,4	1	2
0,25	1	2,5	5
0,5	2	5	10
0,7	2,8	7	14

Aufgabe 7

a) $p = \frac{1}{6}$, n = 60: E(X) = 10 und P(7 ≤ X ≤ 13) ≈ 77,67 %.

b)	n = 4		n = 20		n = 60		n = 100	
	V(X)	σ(X)	V(X)	σ(X)	V(X)	σ(X)	V(X)	σ(X)
p = 0,1	0,36	0,6	1,8	1,34	5,4	2,32	9	3
$p = \frac{1}{6}$	0,56	0,75	2,78	1,67	8,33	2,89	13,88	3,73
p = 0,25	0,75	0,87	3,75	1,94	11,25	3,35	18,75	4,33
p = 0,5	1	1	5	2,24	15	3,87	25	5
p= 0.7	0,84	0,92	4,2	2,05	12,6	3,55	21	4,58

c) Programmierung: In MENU 4 (Tabellenkalkulation) kann man oben zunächst sechs Felder angeben für die Parameter n, p, q, E(X), V(X) und σ(X). Dann gibt man in B1 und B2 Werte für n und p ein. Dann werden über die SEQ Funktion (EDIT-Menu) die Zahlen 0, 1, 2, …, 100 1die dann über FILL (EDIT-Menu) in die Zellen B5 bis B105 kopiert werden. Analoges erfolgt für die Zellen C5 bis C105 und D5 bis D105. Dabei ist bei den Zellen B1 und B2 auf das Zeichen $ zu achten, das vor der Ziffer verhindert, dass beim Nach-unten-Kopieren (FILL) der Zellenbezug fest bleibt. Dann können die Formeln für den Erwartungswert und die Varianz sowie die Standardabweichung programmiert werden.



Beobachtungen: Multipliziert man q mit E(X) ergibt sich die Varianz V(X). Ebenso ergibt sich der Erwartungswert E(X) durch das Produkt von n und p.

Bemerkung: Wenn (n + 1) · p ganzzahlig ist (z. B. n = 9, p = 0,4; μ = 3,6), nimmt die Verteilung an den beiden benachbarten Stellen (n + 1) · p – 1 = μ + p – 1 (= 3) und (n + 1) · p = μ + p (= 4) sein Maximum (= 0,2508) an. Bei nicht ganzzahligem (n + 1) · p (z. B. n = 10, p = 0,4, μ = 4) liegt das einzige Maximum beim größten Wert für k unterhalb von (n + 1) · p = μ + p = 4,4 (also bei 4).

4 Sigma-Regeln

Übungsaufgaben

a)

$\mu \approx 16{,}67$, $V(X) \approx 13{,}89 > 9$, $\sigma \approx 3{,}73$. Es gilt $\mu - 2\sigma \approx 9{,}21$ und $\mu + 2\sigma \approx 24{,}12$. Im Intervall [10, 24] liegen mindestens 95,4 % der möglichen Ausgänge, da $P(10 \le X \le 24) \approx 0{,}957$. Daher würden 10 Einser nicht auf eine signifikante Abweichung hindeuten, während 25 Treffer bei 100 Versuchen auf eine andere Trefferwahrscheinlichkeit hindeuten. Es gilt: $\mu - 2{,}58\sigma \approx 7{,}04$ und $\mu + 2{,}58\sigma \approx 26{,}29$. Daher liegt eine Trefferzahl mit einer Wahrscheinlichkeit von 99 % im Intervall [8; 26].

b)

Sei X die Anzahl der Gewinnlose in einer Stichprobe von $n = 60$. Der Erwartungswert von X beträgt $\mu = 20$. Die Standardabweichung σ beträgt ungefähr $3{,}65 > 3$. Die 2σ-Umgebung lautet $12{,}7 \le X \le 27{,}3$. 95,5 % aller Ausgänge liegen in dieser Umgebung, 4,5 % außerhalb. Das beobachtete Testergebnis liegt linksseitig von der 2σ-Umgebung. Die Wahrscheinlichkeit dafür beträgt nur ca. 2,25 % wegen der nahezu vorliegenden Symmetrie des Histogramms einer Binomialverteilung für große n. Beurteilung: Die Wahrscheinlich ist die Werbung falsch. Der Anteil der Gewinnlose ist vermutlich deutlich geringer als ein Drittel.

c)

Man berechnet zunächst $\mu = 50$ und $\sigma = 5 > 3$. Die Grenzen für das 95 %-Sicherheitsintervall lauten dann $\mu - 1{,}96\sigma = 40{,}2 \le X \le 59{,}8 = \mu + 1{,}96\sigma$. Zum Herausfinden der richtigen Grenzen (die Sigma-Regeln geben nur gerundete Werte an und gelten umso besser je höher n ist) muss man nun verschiedene Möglichkeiten betrachten: $P(41 \le X \le 59) \approx 94{,}3$ % liefert einen zu kleinen Wert. Das Intervall [40; 60] liefert wegen $P(40 \le X \le 60) \approx 96{,}48$ % eine fast 96,5 %-Vorhersagewahrscheinlichkeit, die mindestens 95 % beträgt (nach außen abrunden).

d)

X zählt die Anzahl der gelesenen Exemplare. Man nimmt an, dass X binomialverteilt ist mit $n = 5000$, $p = 0{,}2$, $\mu = 1000$ und $\sigma \approx 28{,}28$. $P(X \le 980) \approx 24{,}6$ %, $P(X \ge 1100) \approx 0{,}025$ %. Mit ca. 95,4 %-iger Wahrscheinlichkeit liegt die Anzahl der ungelesenen Exemplare im Intervall $\mu - 2\sigma \approx 943{,}44 \le X \le 1056{,}56 \approx \mu + 2\sigma$. Das Intervall [944; 1056] liefert die gesuchte Lösung.

5 Testen von Hypothesen mittels Binomialverteilung

Aufgabe 1

b)

Es handelt sich um ein Bernoulli-Experiment unter der Annahme, dass alle Testdurchgänge voneinander unabhängig sind mit jeweils gleicher Trefferwahrscheinlichkeit. Es darf also kein Lerneffekt vorliegen (z. B. durch Ergebnisrückmeldung).

c)

Die Variablen lauten: $n = 40$ und $p = 0{,}5$ (da man zunächst davon ausgehen muss, dass geraten wird). Für den Erwartungswert gilt $\mu = 40 \cdot 0{,}5 = 20$ und die Standardabweichung $\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{\mu \cdot q} = \sqrt{20 \cdot 0{,}5} = \sqrt{10} \approx 3{,}16$. Die Varianz beträgt 10.

Aufgabe 2

a)

$P(X \geq 23) = 1 - P(X \leq 22) \approx 21{,}48\%$

$P(X \geq 27) = 1 - P(X \leq 22) \approx 1{,}924\%$

$P(X \geq 30) = 1 - P(X \leq 22) \approx 0{,}001\%$

b)

Raten kann aufgrund des relativ geringen Stichprobenumfangs nicht ausgeschlossen werden.

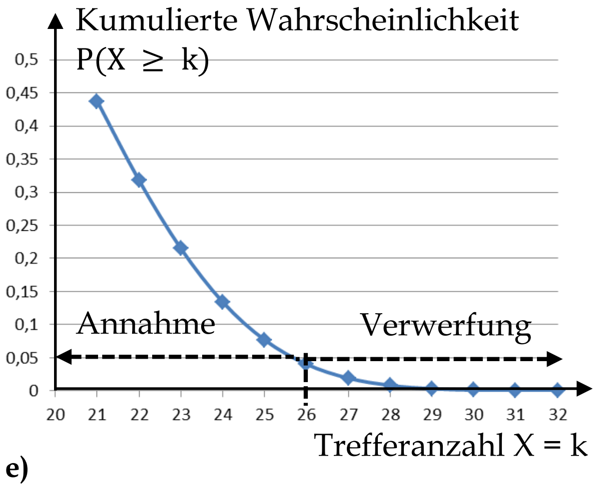
c) und d)

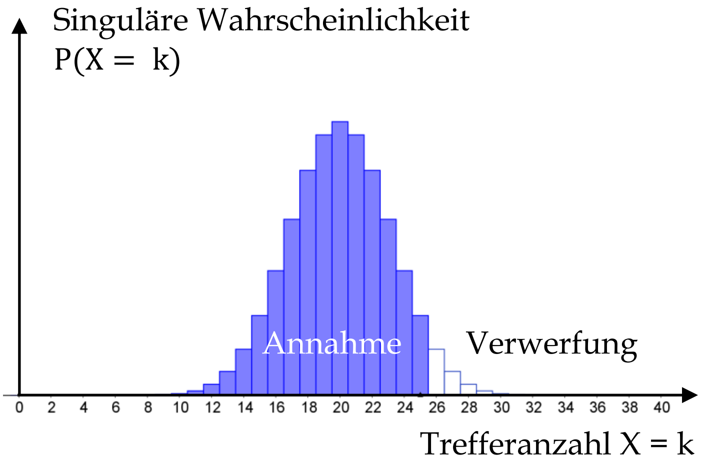
Annahmebereich: $k \leq 25$; Verwerfungsbereich: $k \geq 26$

k	21	22	23	24	**25**	26	27	28	29	30	31
P(x = k)	0,119	0,103	0,081	0,057	**0,037**	0,021	0,011	0,005	0,002	0,001	0,000
P(X ≤ k)	0,682	0,785	0,866	0,923	**0,960**	0,981	0,992	0,997	0,999	0,999	0,999
P(X ≥ k)	0,437	0,318	0,215	0,134	**0,077**	0,040	0,019	0,008	0,003	0,001	0,000

Alternativ können die Sigmaregeln verwendet werden, da $\sigma > 3$: $k \geq 20 + 1{,}64 \cdot 3{,}1622 \approx 25{,}18$. Runden: Es werden sinnvollerweise nur natürliche Zahlen als Werte betrachtet, so dass wir im rechtsseitigen Fall nach oben aufrunden (im linkseitigen Fall würde nach unten abgerundet): Der Verwerfungsbereich liegt bei $k \geq 26$.

Entscheidungsregel: Bestimmt die Versuchsperson nun 26 oder mehr Farben der Gummibärchen richtig, so wird H_0 verworfen und die Alternativhypothese „Die Person ist besser als Raten“ angenommen (jeweils auf einem Signifikanzniveau von 5%). Anderenfalls wird davon ausgegangen, dass die Person nur geraten hat.





e)

Aussage	**Wahr**	**Falsch**	**Begründung**
1. Es ist eindeutig bewiesen, dass die Nullhypothese falsch ist.	☐	■	Man kennt die tatsächliche Trefferwahrscheinlichkeit nicht.
2. Die Wahrscheinlichkeit des Zutreffens der Nullhypothese ist gefunden worden.	☐	■	Man kennt die tatsächliche Trefferwahrscheinlichkeit nicht.
3. Es ist eindeutig bewiesen, mit der die Alternativhypothese wahr ist.	☐	■	Man kennt die tatsächliche Trefferwahrscheinlichkeit nicht.
4. Man kann nun die Wahrscheinlichkeit ableiten, dass die Alternativhypothese richtig ist.	☐	■	Man kennt die tatsächliche Trefferwahrscheinlichkeit nicht.
5. Man geht nun davon aus, dass die Nullhypothese falsch ist, ist sich aber bewusst, dass man bei diesem Verfahren eine wahre Nullhypothese mit einer Wahrscheinlichkeit von maximal 5% fälschlich verwirft.	■	☐	Die Verwerfung der Nullhypothese bedeutet noch nicht, dass sie falsch ist, sondern nur, dass sie zu 5% fälschlich verworfen wurde, falls sie wahr ist.

Aussage	**Wahr**	**Falsch**	**Begründung**
6. Es ist eindeutig bewiesen, dass die Nullhypothese richtig ist	☐	■	Man kennt die tatsächliche Trefferwahrscheinlichkeit nicht.
7. Man kann nun die Wahrscheinlichkeit ableiten, dass die Nullhypothese richtig ist.	☐	■	Man kennt die tatsächliche Trefferwahrscheinlichkeit nicht.
8. Es ist eindeutig bewiesen, dass die Alternativhypothese falsch ist.	☐	■	Man kennt die tatsächliche Trefferwahrscheinlichkeit nicht.
9. Die Wahrscheinlichkeit des Zutreffens der Alternativhypothese ist gefunden worden.	☐	■	Man kennt die tatsächliche Trefferwahrscheinlichkeit nicht.
10. Man geht nun davon aus, dass die Nullhypothese richtig ist, ist sich aber bewusst, dass man bei diesem Verfahren in dem Fall, in dem die Nullhypothese falsch ist, diese mit einer Wahrscheinlichkeit von 5% dennoch fälschlich annimmt	☐	■	Der zweite Teil der Aussage ist falsch, da man die tatsächliche Trefferwahrscheinlichkeit nicht kennt und damit auch nicht die Wahrscheinlichkeit eine falsche Nullhypothese fälschlich anzunehmen.

f)

Stichprobenumfang n	**Kleinstes b mit $P_n(X \leq b) \geq 0,95$**	**Fehler 1. Art:** $P_n(X \geq b)$
10	8	$P_{10}(X \geq 9) \approx 0{,}0107$
20	14	$P_{20}(X \geq 15) \approx 0{,}0207$
40	25	$P_{40}(X \geq 26) \approx 0{,}0403$
50	31	$P_{50}(X \geq 32) \approx 0{,}0325$
100	58	$P_{100}(X \geq 59) \approx 0{,}0443$

Bei größerem Stichprobenumfang nimmt der Annahmebereich bei gleichem p zu. Die Irrtumswahrscheinlichkeit liegt stets unter 5%.

g)

Stichprobenumfang n	**b**	$\mathbf{p_{neu}}$	**Fehler 2. Art:** $P_{p_{neu}}(X \leq b)$
20	14	0,60	87,44%
		0,70	58,36%
		0,80	19,58%
40	25	0,60	68,26%
		0,70	19,26%
		0,80	0,79%
100	58	0,60	37,75%
		0,70	0,72%
		0,80	0,00004%

Durch einen höheren Stichprobenumfang kann man die Wahrscheinlichkeit für den Fehler der 2. Art verringern.

h)

Wenn man den Annahmebereich von H_0 vergrößert, um den Fehler der 1. Art zu verkleinern, wird die Wahrscheinlichkeit für den Fehler der 2. Art vergrößert. Wenn man dann zusätzlich noch den Stichprobenumfang erhöht, kann man gleichzeitig die Wahrscheinlichkeiten für den Fehler der 1. und 2. Art verringern. Für eine wahre Fähigkeit, Gummibärchen an der Farbe zu erkennen, die nur leicht über der Ratewahrscheinlichkeit von 0,5 liegt, ist also der Fehler zweiter Art sehr hoch.

i)

n = 40; p = 0,5	$\alpha = 20\%$	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 1\%$
Annahmebereich (Mit kumulierter Wahrscheinlichkeit)	Kleinstes k mit $P(X \leq k) \geq 0{,}80$ ist 23, also: $k \leq 23$	Kleinstes k mit $P(X \leq k) \geq 0{,}90$ ist 24, also: $k \leq 24$	Kleinstes k mit $P(X \leq k) \geq 0{,}95$ ist 25, also: $k \leq 25$	Kleinstes k mit $P(X \leq k) \geq 0{,}99$ ist 27, also: $k \leq 27$
Mit Sigmaregeln (schlechte Näherungen, da $\sigma = 3{,}16$ nur knapp über 3 liegt)	$k=20+0{,}84 \cdot 3{,}16 \approx 22{,}65$	$k=20+1{,}28 \cdot 3{,}16 \approx 24{,}04$	$k=20+1{,}64 \cdot 3{,}16 \approx 25{,}18$	$k=20+2{,}33 \cdot 3{,}16 \approx 27{,}36$
Verwerfungsbereich	$k \geq 24$	$k \geq 25$	$k \geq 26$	$k \geq 28$

Aufgabe 3

a)

Es handelt sich um ein Bernoulli-Experiment. Da angenommen wird, dass alle Testdurchgänge voneinander unabhängig sind mit jeweils gleicher Trefferwahrscheinlichkeit p = 0,65 (Lose werden immer wieder aufgefüllt). Der Stichprobenumfang beträgt n = 80 und das Signifikanzniveau an jeder Seite 0,5%.

b)

Die Nullhypothese lautet H_0: p = 0,65. Wir nehmen also bei einem Signifikanzniveau von 1% an, dass 65% der Lose Nieten sind.

c)

Die **Gegenhypothese H_A** ist keine Festlegung auf einen bestimmten Anteil der Nieten. Sie wird nur angenommen, wenn das Ergebnis der Stichprobe mit einer Wahrscheinlichkeit von 99% darauf schließen lässt, dass die Nullhypothese nicht zutrifft. Also: H_A: p ≠ 0,65 bei einem Signifikanzniveau von 1%.

d)

(1) X: Anzahl der Nieten; Stichprobenumfang n = 80; Nietenwahrscheinlichkeit p = 65%; Erwartungswert μ = 0,65 · 80 = 52; Standardabweichung $\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{80 \cdot 0{,}65 \cdot\ 0{,}35} \approx 4{,}27$; Signifikanzniveau α = 1%.

(2)	…	39	40	**41**	42	43	…	62	**63**	64	65	…
P(x = k)	…	0,0011	0,0020	**0,0037**	0,064	0,0104	…	0,0056	**0,0029**	0,0015	0,0007	…
P(X ≤ k)	…	0,0021	0,0041	**0,0078**	0,0142	0,0247	…	0,9945	**0,9974**	0,9989	0,9996	…
P(X ≥ k)	…	0,9990	0,9979	**0,9959**	0,9922	0,9858	…	0,0111	**0,0055**	0,0026	0,0004	…

e)

(1) Kleinste Zahl a, bei der 0,005 überschritten wird: a = 41.

(2) Kleinste Zahl b, bei der 0,995 überschritten wird: b = 63.

(3) **Annahmebereich**: 41 ≤ k ≤ 63; **Verwerfungsbereich**: k ≤ 40 und k ≥ 63; **Entscheidungsregel**: Das bedeutet, dass für den Fall, dass der Mathematik-LK zwischen 41 und 63 Nieten nachgewiesen hat, die Annahme, dass der tatsächliche Anteil der Nieten 65% beträgt, akzeptiert wird. In allen anderen Fällen geht der Mathe-LK davon aus, dass der tatsächliche Anteil der Nieten nicht 65% beträgt.

f)

Bei einem Signifikanzniveau von 10% wäre der Annahmebereich (mit a_1= [45;59], wie entsprechend nachzurechnen ist) deutlich kleiner. Das bedeutet, dass die Irrtumswahrscheinlichkeit entsprechend größer wäre. Dem Hersteller gegenüber können die Schüler des Mathe-LKs bei einem Ergebnis im Verwerfungsbereich auf dem Signifikanzniveau von 1% entsprechend sicherer sein können, dass der Hersteller einen Produktionsfehler gemacht hat, als bei einem Signifikanzniveau von 10%. Allerdings fällt dieser auch erst „später“ auf, d. h., dass bei einer tatsächlichen Wahrscheinlichkeit von p ≠ 0,65 länger die Nullhypothese noch (fälschlicherweise) akzeptiert wird, der Produktionsfehler also „weniger wahrscheinlich entdeckt“ wird.

g)

Dieses Vorgehen ist nicht sinnvoll, sofern der Ansatz gewählt wird, dass neutral aufgrund einer vorher aufgestellten Hypothese deren Gültigkeit überprüft wird. Der Test ist nur nachvollziehbar konzipiert, wenn das Signifikanzniveau vorher angegeben wird. Zudem sagt der Test ja auch nicht aus, ob eine Hypothese wahr ist oder falsch, sondern führt zu einer Entscheidung für oder gegen die Nullhypothese. Wird das Signifikanzniveau im Nachhinein festgelegt, so kann genau diese Entscheidung im Sinne des Testers beeinflusst werden. Bemerkung: Allerdings ist es sinnvoll, wenn der Mathe-LK anhand möglicher Signifikanzniveaus die Bedeutung und die Abhängigkeit der Fehler erster und zweiter Art verdeutlichen und erkennen möchte ☺.

h)

(1) Der Fehler erster Art gibt die Wahrscheinlichkeit dafür an, dass H_0 fälschlicherweise verworfen wird, d. h., dass weniger als 41 oder mehr als 63 Nieten gezogen werden, obwohl tatsächlich 65% Nieten enthalten sind.

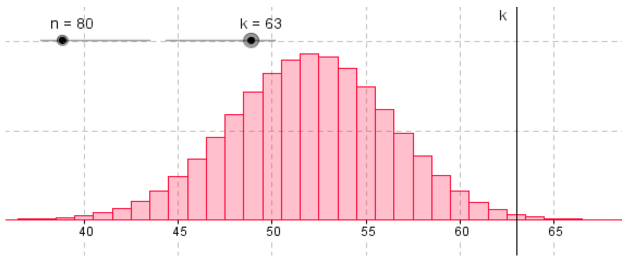
(2) $\alpha = P(X \leq 40) + P(X \geq 64) = P(X \leq 40) + 1 - P(X \leq 63) = 0{,}0041 + 0{,}0026 = 0{,}0067 = 0{,}67\%$
Anschaulich ist der Fehler erster Art die Summe der beiden rot markierten Werte in der obigen Tabelle.

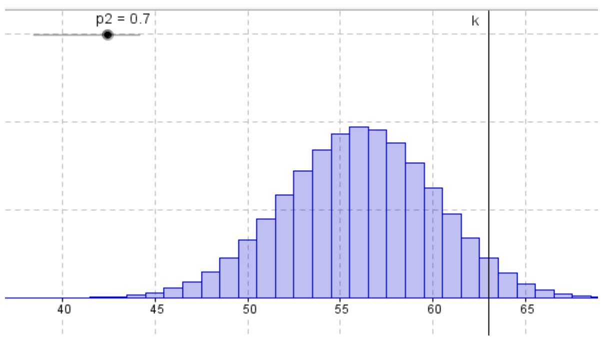
i)

(1) Der Fehler zweiter Art kann nicht bestimmt werden, da in dieser Aufgabe die „echte“, also die tatsächliche Wahrscheinlichkeit unbekannt ist. Im Zusammenhang gibt der Fehler zweiter Art die Wahrscheinlichkeit an, mit der von einem 65%-Nietenanteil ausgegangen wird, obwohl tatsächlich mehr oder weniger Nieten enthalten sind.

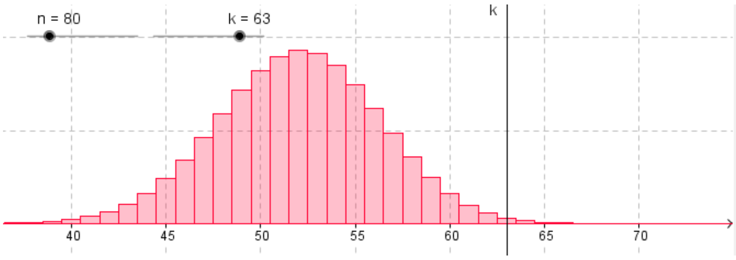
(2) Um den Fehler zweiter Art zu berechnen, betrachtet man den Annahmebereich der Nullhypothese unter der Voraussetzung, dass die Alternativhypothese gilt. Der Fehler zweiter Art ist die Wahrscheinlichkeit dafür, dass ein Testergebnis in den Annahmebereich der Nullhypothese fällt, obwohl die Alternativhypothese gilt.

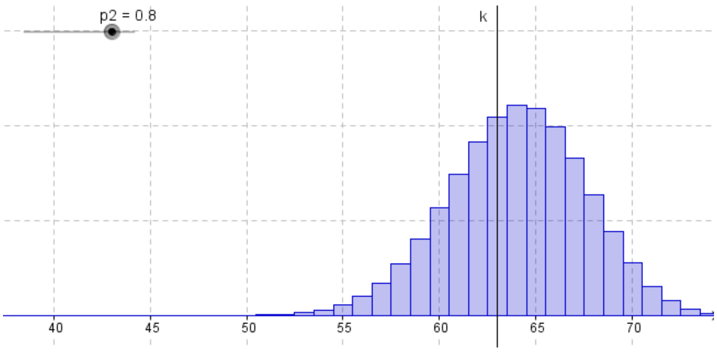
(3) Auf der „rechten Seite“ heißt das, dass der Fehler zweiter Art die Wahrscheinlichkeit dafür ist, dass unter der Annahme, dass p = 0,7 gilt, das Ergebnis im Intervall a = [41;63] liegt. Das berechnet sich mit Hilfe der kumulierten Binomialverteilung: $P_{0,7}(41 \leq X \leq 63) = 0{,}9697$.





(4) Der Fehler zweiter Art wird deutlich geringer, da deutlich weniger Ergebnisse im Annahmebereich von H_0 liegen.





j)

(1) $\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{80 \cdot 0{,}65 \cdot\ 0{,}35} \approx 4{,}27$
(2) a =2,58 (da das Signifikanzniveau 1% gewählt wurde).
(3) a = [μ – 2,58σ; μ + 2,58σ] = [52 – 2,58 ·4,27; 52 + 2,58 ·4,27] = [40,98; 63,01] ≈ [41; 63].

Exkurs zur Rundung beim Rechnen mit Sigmaregeln: Im Annahmebereich liegen mindestens 95% der Ergebnisse, im Verwerfungsbereich maximal 5% der Ergebnisse. Damit dies mit Sicherheit noch gilt (insbesondere bei kleinen σ-Werten kann es sonst dazu kommen, dass nicht mehr 95% der Ergebnisse im Annahmebereich liegen), muss der Annahmebereich rundungstechnisch „vergrößert“ werden, damit im Zweifelsfall noch mehr Ergebnisse im Annahmebereich liegen. Das bedeutet hier dann a= [40;64], um sicher zu sein.[8] Allerdings kann die Rundungsfrage differenzierter betrachtet werden: Es gibt theoretisch vier Rundungsmöglichkeiten. Von jedem dieser Intervalle kann man P(A) berechnen:

- a_1 = [40;63], $P(40 \leq X \leq 63) = 0{,}995372$
- a_2 = [41;63], $P(41 \leq X \leq 63) = 0{,}993327$
- a_3 = [40;64], $P(40 \leq X \leq 64) = 0{,}996825$
- a_4 = [41;64], $P(41 \leq X \leq 64) = 0{,}99478$

Es ist nun derjenige Annahmebereich zu wählen, der den geringsten Abstand zu 0,99 hat. In diesem Fall ist die beste Rundung somit a_2 = [41;63]. Dies muss nicht immer dem Annahmebereich entsprechen, der sich durch „nach außen runden“ (hier: a_3 = [40;64]) oder durch mathematische Rundungsregeln (die ohnehin hier nicht sinnvoll anzuwenden sind) ergibt. Die Auswahl des Annahmebereichs hat dann auch Auswirkungen auf die Fehler erster und zweiter Art.

[8] In den Abitur-Modelllösungen wird das nicht differenzierter betrachtet. Mit Hilfe des GTR ist dies aber ohne größeren Aufwand möglich.

Aufgabe 4

a) und b)

Testarchitektur, Binomialverteilungen: X = Anzahl der Verspätungen; n = 100; p = 0,40; Signifikanzniveau α = 5%. Daher gilt: $\mu = n \cdot p = 40$; $\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{100 \cdot 0{,}40 \cdot 0{,}60} \approx 4{,}90$.

Hypothesenbildung: Es sei X die Anzahl der Verspätungen. Die Bürgerinitiative ist überzeugt von der Richtigkeit ihrer Behauptung und wird sich nur vom Gegenteil überzeugen lassen, wenn sehr kleine Werte von X auftreten. Daher testet sie nicht das Gegenteil ihrer Behauptung. Es geht also um einen linksseitigen Test. Die **Nullhypothese H_0** ist **$p \geq 0{,}4$** (man könnte auch p = 0,4 wählen, da bei Werten über 40 % die Nullhypothese ebenfalls erfüllt wäre). Die **Alternativhypothese H_A** ist **$p < 0{,}4$** (Das Eintreten wäre aus der Sicht der Bahn positiv).

c)

k	…	29	30	31	32	33	34	35	36	37	…
P(X ≤ k)	...	0,0148	0,0248	**0,0398**	0,0615	0,0913	0,1303	0,1795	0,2386	0,3068	…

d)

Signifikanz und Relevanz, Annahme und Verwerfungsbereich: Beim Ablesen der Tabelle wird P(X ≤ k) das erste Mal für k = 32 größer als 5%. Der **Annahmebereich** ist somit A = [32;100] und der **Verwerfungsbereich** V = [0;32]. **Entscheidungsregel**: Im Anwendungszusammenhang bedeutet das, dass ab 32 Verspätungen mit einer Wahrscheinlichkeit von über 95% davon auszugehen ist, dass die tatsächliche Verspätungsquote (mindestens) 40% beträgt, da mehr als 95% der möglichen Ergebnisse in diesem Intervall liegen. Bei 32 Verspätungen oder mehr wird also die Nullhypothese angenommen, bei weniger als 32 Verspätungen wird sie verworfen.

e)

Fehler 1. Art beim Testen: Der Fehler erster Art beträgt $\alpha = P(X \leq 31) \approx$ 3,98% (siehe Tabelle oben). Das bedeutet, dass mit 4% Wahrscheinlichkeit die Nullhypothese (die Bahn hat mindestens 40% Verspätungen) abgelehnt wird und fälschlicherweise (wenn die Verspätungsquote mindestens 40 % beträgt) von einem geringeren Anteil an Verspätungen ausgegangen wird.

f)

Der **Fehler zweiter Art** (unter der Annahme, dass die tatsächliche Verspätungswahrscheinlichkeit 30% beträgt) beträgt $\beta = P_{0,3}(32 \leq X \leq 100) = 1 - P_{0,3}(X \leq 31) \approx 36{,}67\%$. Das bedeutet, dass der Ausgang der Stichprobe im Annahmebereich liegt, obwohl die tatsächliche Wahrscheinlichkeit für Verspätungen geringer ist. Aus Sicht der Bürgerbewegung ist der recht große Fehler zweiter Art natürlich unproblematisch, da die These nachgewiesen werden soll, dass die Verspätungsquote (mindestens) 40% beträgt. Das würde der VRR anders bewerten, da das Ergebnis der Untersuchung in eben mehr als einem Drittel der Untersuchungsergebnisse schlechter ausfallen würde als das Angebot der Bahn eben ist.

g)

Sigma-Regeln: $\mu = n \cdot p = 40$; $\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{100 \cdot 0{,}40 \cdot 0{,}60} \approx 4{,}90$; a = 1,64 (da das Signifikanzniveau 5% gewählt wurde).

$A = [\mu - 1{,}64 \cdot \sigma\,;\; n] = [40 - 1{,}64 \cdot 4{,}899\,;100] = [31{,}96\,;100] \approx [32\,;100].$

h)

Der VRR wird den (mit ca. 36%) aus ihrer Sicht zu hohen Fehler zweiter Art als einen Nachteil sehen. Ein **rechtsseitiger** Test kommt zu einem Annahmebereich, der besagt, dass bei einem Untersuchungsergebnis zwischen 0 und 48 Verspätungen mit einer Wahrscheinlichkeit von 95% die Verspätungsquote höchstens 40% beträgt.

k	…	45	46	47	48	49	50	51	52	…
P(X ≤ k)	...	0,8689	0,9070	0,9362	0,**9577**	0,9729	0,9832	0,9900	0,9942	…
P(X ≥ k)	…				0,0638	**0,0423**	0,0271			

Bezüglich möglicher Werbeaussagen und Imagefragen ist diese Vorgehensweise natürlich aus Sicht der Bahn besser. Hinzu kommt ein für sie günstiger recht großer Fehler zweiter Art. Denn in ca. 10% der Fälle wird die für die Bahn günstigere Aussage fälschlicherweise beibehalten. Die Bürgerbewegung hingegen kann nur ca. 4% Fehler erster Art „erhoffen", nachdem die Nullhypothese fälschlicherweise verworfen wird. Außerdem ist der Annahmebereich aus Sicht der Bürgerbewegung (die ja von mindestens 40% ausgeht) sehr groß, da es einen hohen Überschneidungsbereich der beiden Annahmebereiche gibt: A = [32;48]. Solange dies der Fall ist, ist die Deutung des Untersuchungsergebnisses stark von der Testarchitektur abhängig und weniger vom Untersuchungsergebnis. In keinem Fall lässt sich mit Hilfe der Untersuchung sagen, wie groß die tatsächliche Verspätungsquote der Bahn ist.

Aufgabe 5: Münzwurf

a) Bei jedem Wurf der beiden Münzen wird als Treffer das Ergebnis „beide Münzen zeigen Zahl" betrachtet. Die einzelnen Würfe sind unabhängig voneinander. Somit ist es eine Bernoulli-Kette. Die Länge ist n = 50 und die Trefferwahrscheinlichkeit ist p = 0,25.
b) P(X = 12) = 12,94%; $P(X \leq 12) = 51{,}10\%$;
c) $P(X \geq 12) = 1 - P(X \leq 11) = 61{,}84\%$; $P(15 \leq X \leq 30) = 25{,}19\%$
d) σ = 50 · 0,25 = 12,5. Wenn man die Bernoulli-Kette sehr oft durchführt, sind im Durchschnitt auf lange Sicht 12,5 Treffer zu erwarten.
e) σ = 3,0619
f) Das 2σ-Intervall: [7;18].

Aufgabe 6: Kirschkerne

a) n = 100, p = 0,02; $P(X \geq 1) = 86{,}74\%$
b) μ = 2
c) p = 0,015, k = 0, n gesucht. Er darf höchstens 11 Kirschen nehmen.
d) n = 120, k = 0, p gesucht. Die Wahrscheinlichkeit darf höchstens 0,2229% betragen.

Aufgabe 7: Brot auf die Marmeladenseite

a) H_0: p = 0,5; H_A: p ≠ 0,5; X ist binomialverteilt mit Parametern n = 25 und p = 0,5.
b) Annahmebereich [8; 17], d.h. wenn das Brot zwischen 8-und 17-mal auf die Marmeladenseite fällt, behält Katja die Nullhypothese bei.
c) P(X < 8) + P(X > 17) = 0,0433

d) $P_{0,7}(8 \leq X \leq 17) = 0{,}4881$

Aufgabe 8: Umfrage zur Stadthalle

a) Linksseitiger Test. Die Nullhypothese, dass 75 % für den Bau sind, soll verworfen werden, wenn es in der Umfrage nur wenig Befürworter gibt.
H_0: $p = 0{,}75$; H_A: $p < 0{,}75$; Annahmebereich: [68; 100]
b) Bei höchstens 67 Befürwortern kann die Redaktion die Schlagzeile drucken.
c) $P(X \leq 67) = 0{,}0446$
d) $P_{0,6}(X \geq 68) = 0{,}0615$

6 Kontrollaufgaben

Aufgaben ohne Zuhilfenahme von Hilfsmitteln

Aufgabe 1

a) Man rechne nach: $\mu(\text{Mathematik}) = \frac{1}{24}(1 \cdot 6 + 2 \cdot 4 + 3 \cdot 2 + 4 \cdot 2 + 5 \cdot 4 + 6 \cdot 6) = \frac{84}{24} = \frac{21}{8} = 3{,}5$
$\mu(\text{Englisch}) = \frac{1}{24}(1 \cdot 2 + 2 \cdot 4 + 3 \cdot 6 + 4 \cdot 6 + 5 \cdot 4 + 6 \cdot 2) = \frac{84}{24} = \frac{21}{8} = 3{,}5$. Im Fach Englisch liegen offenbar mehr Klassenarbeiten in der Nähe des Erwartungswertes. Dort ist die Streuung geringer. Daher sind (1) und (3) richtig.

b) (1) In diesem Fall müsste wegen der Symmetrie $P(X = 0) = P(X = 2) = 0{,}5 \cdot 0{,}5 = 0{,}25$ sein.

(2) $\mu = 0 \cdot 0{,}32 + 1 \cdot 0{,}36 + 2 \cdot 0{,}32 = 1$; $V = (0-1)^2 \cdot 0{,}32 + (1-1)^2 \cdot 0{,}36 + (2-1)^2 \cdot 0{,}32 = 0{,}64$;
Daher beträgt die Standardabweichung $\sigma = = \sqrt{V} = \sqrt{0{,}64} = 0{,}8$.

Aufgabe 2

a) X: Anzahl der blauen Kugeln; $p = 40\,\%$, $k \le 1$
$P(X \le 1) = \binom{3}{0} 0{,}4^0 \cdot 0{,}60^3 + \binom{3}{1} 0{,}4^1 \cdot 0{,}60^2 = 0{,}60^3 + 3 \cdot 0{,}4 \cdot 0{,}36$
$= 0{,}60^3 + 3 \cdot 0{,}4 \cdot 0{,}36 = (0{,}6 + 1{,}2) \cdot 0{,}36 = 0{,}216 + 0{,}432 = 0{,}648$

b) $P(X = k) = \binom{4}{k} 0{,}4^k \cdot 0{,}60^{4-k}$ für $0 \le k \le n = 4$.

Aufgabe 3

a)	M	W	
60+	$0{,}8 \cdot 37{,}5\,\% = 30\,\%$	40 %	**70 %**
unter 60	7,5 %	22,5 %	30 %
	37,5 %	62,5 %	

P(W und 60+) = 40 %.

b) Drei Achtel (= 37,5 %) der Reisegruppe sind Männer, fünf Achtel (62,5 %) sind Frauen. Daher müssen zwei Achtel der Gesamtgruppe der Personenzahl 10 entsprechen. Insgesamt hat die Reisegruppe 40 Personen (25 Frauen und 15 Männer).

Aufgabe 4

a) Nur (2) ist falsch. Alle anderen Terme sind gleich, da $\binom{20}{3} = \binom{20}{17}$ und $0{,}5^3 \cdot 0{,}5^{17} = 0{,}5^{20}$ sowie $\binom{20}{3} = \frac{20 \cdot 19 \cdot 18}{3!}$.

b) A gehört zu T_3, B gehört zu T_1, C gehört zu T_2

Aufgabe 5

a) Der Erwartungswert von X beträgt $10 \cdot 0{,}4 = 4$ und von Y $10 \cdot 0{,}6 = 6$. Da der Erwartungswert in der Nähe der Trefferzahl mit der größten Wahrscheinlichkeit liegt, handelt es sich um die Verteilung von X. Ferner gilt: $P(X \neq 5) = 1 - P(X = 5) = 1 - 0{,}20 = 0{,}80$.

b) $P(X = k) = \binom{10}{k} 0{,}4^k \cdot 0{,}60^{10-k} = \binom{10}{10-k} 0{,}60^{10-k} \cdot 0{,}4^{10-(10-k)} = P(Y = 10 - k)$.
Alternativ kann auch argumentiert werden, dass Y die komplementäre Zufallsgröße zu X ist (Trefferwahrscheinlichkeit von X ist Nichttrefferwahrscheinlichkeit von Y).
Ferner gilt: $P(Y = 6) = P(X = 4) = 0{,}25$

Aufgabe 6

a) $\mu = 100 \cdot 0{,}2 = 20;\ \sigma = \sqrt{\mu \cdot (1 - p)} = \sqrt{20 \cdot 0{,}8} = \sqrt{16} = 4$

b) $V = n \cdot p \cdot (1 - p) \Leftrightarrow 9 = 100 \cdot p \cdot (1 - p) \Leftrightarrow \frac{9}{100} = p \cdot (1 - p) \Leftrightarrow p^2 - p + 0{,}09 = 0$
$\Leftrightarrow p = 0{,}5 \pm \sqrt{0{,}25 - 0{,}09} = 0{,}5 \pm \sqrt{0{,}16} = 0{,}5 \pm 0{,}4$. Daher gilt $p = 0{,}9$ oder $p = 0{,}1$ mit den Erwartungswerten $\mu = 90$ oder $\mu = 10$.

c) $\sigma = \sqrt{n \cdot p \cdot (1 - p)}$: Je größer n ist, desto größer ist die Standardabweichung. Sie gibt die Streuung um den Erwartungswert an. Deshalb ist die Streuung um den jeweiligen Erwartungswert für n = 80 am größten und für n = 20 am kleinsten.

Aufgaben mit Zuhilfenahme von Hilfsmitteln

Aufgabe 7

a) Voraussetzungen:

- Genaue eine Antwort ist richtig.
- Nur 2 mögliche Ausgänge sind möglich (richtige Antwort = Treffer, falsche Antwort = Niete),
- Alle Aufgaben sind gleich schwer und hängen nicht voneinander ab.
- Kenntnisstand = Trefferwahrscheinlichkeit p
- Keine Ermüdung
- Zufallsgröße X: Anzahl der richtig beantworteten Fragen

b) (1) $n = 20$, $k \geq 12$ (60 % von 20), $p = 0{,}7 \Rightarrow P(X \geq 12) \approx 88{,}67$ %;
(2) Wegen $\mu = 20 \cdot 0{,}7 = 14$ ist die kumulierte Wahrscheinlichkeit gesucht: $P(9 \leq X \leq 19) \approx 99{,}41$ %.

c) (1) P(1. bis 8. Frage richtig und 9. sowie 10. Frage falsch) $= 0{,}8^8 \cdot 0{,}2^2 \approx 0{,}67$ %.
(2) $n = 10$, $k = 8$, $p = 0{,}8$: $P(X = 8) \approx 30{,}20$ %.

d) Gegeben: $n = 20$, $k \geq 15$, $P_{n=20;p}(X \geq 15) \geq 0{,}9$. **Gesucht**: (Mindest-)Kenntnisstand p.
Betrachte im Graphen-Menu die Funktion f mit $f(x) = P_{n=20;x}(X \geq 15)$ (Eingabe: Y1 = BinomialCd(15, 20, 20, x)) (x: Kenntnisstand) und bringe sie mit der zweiten Funktion g mit $g(x) = 0{,}9$ (Eingabe: Y2 = 0.9) zum Schnitt. Man erhält die Schnittstelle $x \approx 0{,}8341$, an der $P_{n=20;x}(X \geq 15) = 0{,}9$ gilt. Für $x \geq 0{,}8341$ gilt dann $P_{n=20;x}(X \geq 15) \geq 0{,}9$. Damit beträgt der Kenntnisstand von Bewerber C mindestens 83,41 %.

e) Gegeben: $n = 20$, $p = 0{,}9$, maximales P. **Gesucht**: k
Betrachte im Tabellen-Menu die Funktion f mit $f(x) = P_{n=20;p=0{,}9}(X = x)$ (Eingabe: Y1 = BinomialPd(x, 20, 0.9))(x: Anzahl der richtigen Fragen) und suche in der Tabelle den k-Wert mit der größten Wahrscheinlichkeit. Für $k = 18$ wird $P(X = k)$ maximal, denn es gilt: $P(X = 17) \approx 0{,}1901 < P(X = 18) \approx 0{,}2851 < P(X = 19) \approx 0{,}2701$. Alternativ kann mit den Erwartungswert $\mu = 20 \cdot 0{,}9 = 18$ argumentiert werden, beim dem aufgrund der Ganzzahligkeit die größte Wahrscheinlichkeit vorliegt.

f) Gegeben: $k \geq 10$, $p = 0{,}75$, $P_{n;p=0{,}75}(X \geq k) \geq 0{,}95$. **Gesucht**: Anzahl der Fragen n
Betrachte im Tabellen-Menu die Funktion f mit $f(x) = P_{n;p=0{,}75}(X \geq 10)$ (Eingabe: Y1 = BinomialCd(10, x, x, 0.75) (x: Anzahl der Fragen) und schaue, für welches x f (hier als Y1) größer als 0,95 wird. Bei $n = 17$ Fragen ist $P(X \geq 10) \approx 0{,}9597 > 95$ %. Also müssten mindestens 17 Fragen gestellt werden, damit der Proband mindestens 10 Fragen mit einem Kenntnisstand von 75 % richtig beantwortet.

g) Betrachte die Vierfelder-Tafel (oder ein Baumdiagramm) mit den fettgedruckten Eintragungen für gegebenen Informationen: $P(B) = 68$ % haben sich insgesamt bewährt.

	Test bestanden (A)	Durchgefallen ($\overline{A}$)	
Bewerber bewährt sich (B)	$0{,}8 \cdot 0{,}8 = 0{,}64$	0,04	0,68
Bewerber bewährt sich nicht ($\overline{B}$)	0,16	$0{,}8 \cdot 0{,}2 = 0{,}16$	0,32
	0,8 = 80 %	**0,2 = 20 %**	

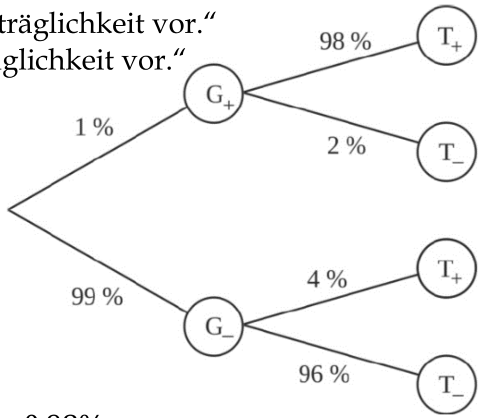
h) Gegeben: $n = 60$, $p = 0{,}6$, $P(X \geq k) \leq 0{,}75$. **Gesucht**: Minimale Anzahl k der richtigen Antworten.
Betrachte im Tabellen-Menu die Funktion f mit $f(x) = P_{n=20;p=0{,}6}(X \geq x)$ (Eingabe: Y1 = BinomialCd(x, 60, 60, 0.6) (x: Anzahl der richtigen Fragen) und suche in der Tabelle den kleinsten k-Wert mit $P(X \geq k) \leq 0{,}75$. Für $k = 33$ gilt $P(X \geq k) \approx 0{,}8221$. Für $k = 34$ gilt $P(X \geq k) \approx 0{,}7464$. Daher ist $k = 34$ der gesuchte Wert.

i) Der Erwartungswert μ beträgt 60 · 0,6 = 36. Die Standardabweichung σ liegt bei $\sqrt{\mu \cdot (1-p)} = \sqrt{36 \cdot 0{,}4} = 1{,}2 \cdot \sqrt{10} \approx 3{,}79 > 3$. Daher lautet die 2σ-Umgebung von 36: [28,41; 43,59]. Betrachte nun die kumulierte Wahrscheinlichkeit $P(29 \leq X \leq 43) \approx 95{,}3$ %. Sollten mehr als 4,7 % der Testergebnisse mit mehr als 43 oder weniger als 29 richtigen Antworten erfolgen, liegt eine signifikante Abweichung vom Erwartungswert vor und der angenommene Schwierigkeitsgrad wäre zu überdenken.

Aufgabe 8

a)

(1) G_+: „Bei der Person liegt eine Glutenunverträglichkeit vor."
G_-: „Bei der Person liegt keine Glutenunverträglichkeit vor."
T_+: „Das Testergebnis ist positiv."
T_-: „Das Testergebnis ist negativ."



(2) $P(A) = P(G_+ \cap T_+) = 0{,}01 \cdot 0{,}98 = 0{,}0098 = 0{,}98\%$
$P(B) = P(T_-) = P(G_+ \cap T_-) + P(G_- \cap T_-) = 0{,}01 \cdot 0{,}02 + 0{,}99 \cdot 0{,}96 = 0{,}9506 = 95{,}06\%$

(3) $P_{T_+}(G_+) = \frac{P(G_+ \cap T_+)}{P(T_+)} = \frac{0{,}0098}{1-P(T_-)} = \frac{0{,}0098}{0{,}0494} \approx 0{,}19840 = 19{,}84\%$.

b)

(1) Das Modell der binomialverteilten Zufallsgröße ist hier geeignet, da es sich um eine diskrete Zufallsgröße handelt und ein Bernoulli-Prozess mit genau zwei möglichen Ergebnissen („krank" und „gesund") vorliegt. Darüber hinaus kann man aufgrund der großen Grundgesamtheit (82 Millionen) und der kleinen Stichprobe (20 000) von einer annähernd gleichbleibenden Wahrscheinlichkeit für die Auswahl einer von Glutenunverträglichkeit betroffenen Person ausgehen.

(2) $P(E_1) = P_{n=20000;p=0{,}01}(X = 190) \approx 2{,}25\%$; $P(E_2) = P_{n=20000;p=0{,}99}(X > 19800) \approx 49{,}05\%$
$P(E_3) = P_{n=20000;p=0{,}01}(240 \leq X \leq 2400) \approx 0{,}31\%$

(3) X ist binomialverteilt mit n = 20000 und p = 0,01. $\mu = E(X) = n \cdot p = 20000 \cdot 0{,}01 = 200$. Die gesuchte Wahrscheinlichkeit beträgt: $1 - P(0{,}9 \cdot 200 \leq X \leq 1{,}1 \cdot 200) \approx 0{,}1450 = 14{,}50$ %.

c)

(1)

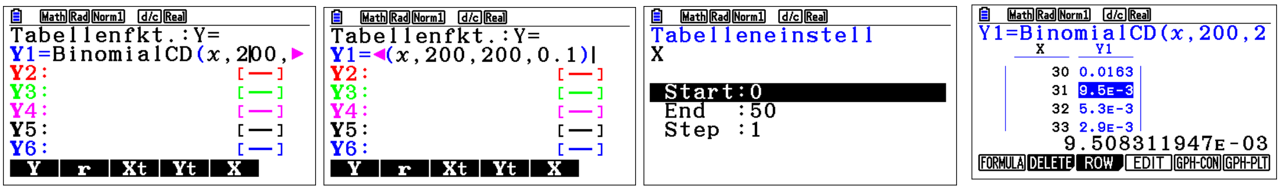
Folgende Fehlentscheidungen können auftreten:

- Obwohl höchstens 10 % der hergestellten Teststreifen unbrauchbar sind, entscheidet man sich aufgrund des Ergebnisses der Kontrolle dafür, das Herstellungsverfahren zu verbessern. (Fehler 1. Art = α-Fehler)

- Obwohl mehr als 10 % der hergestellten Teststreifen unbrauchbar sind, entscheidet man sich aufgrund des Ergebnisses der Kontrolle nicht dafür, das Herstellungsverfahren zu verbessern. (Fehler 2. Art = β-Fehler)

(2) Die Zufallsgröße Y beschreibt die Anzahl der Teststreifen in der Stichprobe, die unbrauchbar sind. Y ist binomialverteilt mit n = 100 und p = 0,1. Die Wahrscheinlichkeit für mindestens 16 unbrauchbare Teststreifen beträgt $P(16 \le Y \le 100) \approx 0{,}0399 = 3{,}99\ \%$.

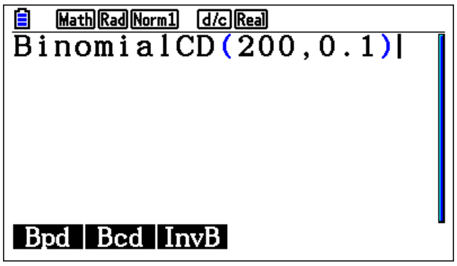
(3) Gegeben: Y: „Anzahl der unbrauchbaren Teststreifen“ ist binomialverteilt mit n = 200 und Trefferwahrscheinlichkeit p = 0,1 sowie $P_{n=200;p=0,1}(Y \ge k) \le 1{,}33\ \%$. **Gesucht**: Mindestzahl k
Man definiert im Tabellenmenu (MENU 7) eine Funktion f mit $f(k) = P_{n=200;p=0,1}(Y \ge k)$ mit der Trefferzahl k. Es gilt $f(30) = P_{n=200;p=0,1}(Y \ge 30) \approx 0{,}016 > 1{,}33\%$ und $f(31) = P_{n=200;p=0,1}(Y \ge 31) \approx 0{,}001 < 1{,}33\%$. Daher gilt für $k \ge 31$ $P_{n=200;p=0,1}(Y \ge k) \le 1{,}33\ \%$. Die Verwendung des GTR ist im Folgenden dokumentiert:

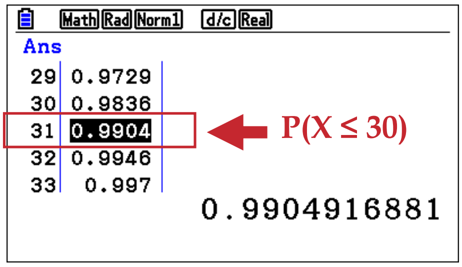


Alternativlösung: Einige Schüler favorisieren die Listenfunktion der kumulierten Wahrscheinlichkeiten in MENU 1. Man müsste die **Ausgangsfrage** umformulieren:

Für welche z der binomialverteilten Größe Y: „Anzahl der unbrauchbaren Teststreifen“ mit p = 0,1 und n = 200 gilt $P(Y \le z) \ge 98{,}67\ \%$?

Für $z \ge 30$ ist $P(Y \le z) \ge 98{,}67\ \%$. Damit kann für alle $k \ge 31$ Folgendes gefolgert werden: $P(Y \ge k) = 1 - P(Y \le k-1) \le 1 - 98{,}67\ \% = 1{,}33\ \%$.





(4) Y ist nun binomialverteilt mit n = 100 und p = 0,18. $P(Y < 16) = P(Y \le 15) \approx 26{,}30\%$.

d) (1) $\bar{x} = \frac{1}{100} \cdot (4 \cdot 15 + 9 \cdot 16 + 10 \cdot 17 + 48 \cdot 18 + 18 \cdot 19 + 11 \cdot 20) = 18$ mg

$$\sigma = \sqrt{\frac{4\cdot(18-15)^2+9\cdot(18-16)^2+10\cdot(18-17)^2+48\cdot(18-18)^2+18\cdot(18-19)^2+11\cdot(18-20)^2}{100}} = 1{,}2\ \text{mg}$$

(2) Das arithmetische Mittel ist konstant geblieben, d. h., die durchschnittliche Indikatormenge pro Teststreifen hat sich nicht verändert. Eine deutliche Veränderung gibt es bei der Standardabweichung, die von 4,3 mg auf 1,2 mg gesunken ist. Eine geringere Streuung bedeutet, dass die Indikatormengen auf den Teststreifen weniger stark von der durchschnittlichen Menge abweichen als vorher. Das Produktionsverfahren ist somit präziser und somit qualitativ besser geworden.

Aufgabe 9

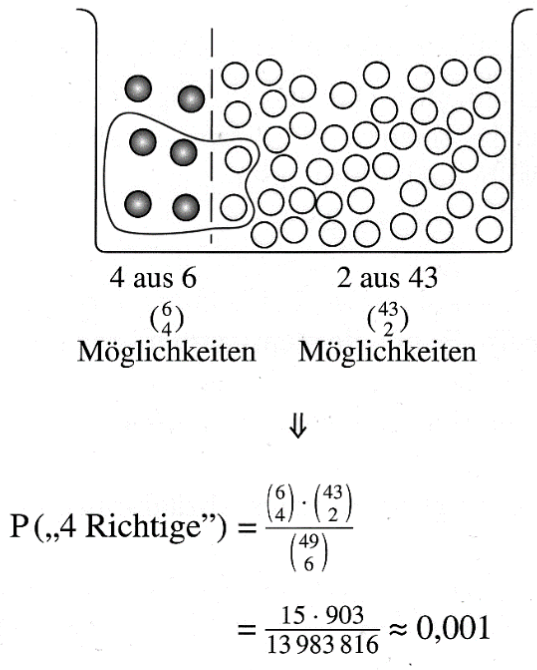
a) Sei X: Anzahl der richtigen Zahlen im „Lotto 6 aus 49“. Es gibt genau $\binom{49}{6}$ unterschiedliche Kombinationen (ohne Beachtung der Reihenfolge). Daher gilt $P(X = 6) = \frac{1}{\binom{49}{6}}$.

Alternativ kann mithilfe der sechs Pfadwahrscheinlichkeiten die Wahrscheinlichkeit für sechs Richtige durch das Produkt $\frac{6}{49} \cdot \frac{5}{48} \cdot \frac{4}{47} \cdot \frac{3}{46} \cdot \frac{2}{45} \cdot \frac{1}{44} = \frac{6!}{49 \cdot 48 \cdot \ldots \cdot 45 \cdot 44} = \frac{1}{\binom{49}{6}}$ berechnet werden.

b) Insgesamt sind wegen a) $\binom{49}{6}$ Tipps möglich. Um festzustellen, wie viele dieser Tipps günstig sind für das Ereignis E: „Vier Richtige“ verwendet man folgende Grundidee: Man denkt sich den Inhalt der Urne in zwei Gruppen: einmal die 6 Richtigen und zum anderen die 43 Nieten.

- Ein für E günstiger Tipp besteht aus 4 Richtigen und 2 Nieten.
- Es gibt $\binom{6}{4} = 15$ Möglichkeiten, aus der Gruppe der 6 Richtigen 4 Richtige auszuwählen.
- Analog gibt es $\binom{43}{2} = 903$ Möglichkeiten aus der Gruppe von 43 Nieten 2 Nieten auszuwählen.
- Daher gibt es insgesamt $\binom{6}{4} \cdot \binom{43}{2}$ Möglichkeiten, vier Treffer mit zwei Nieten zu einem für E günstigen Tipp zu kombinieren.
- Dividiert man diese Zahl durch die Anzahl der möglichen Kombinationen erhält man die gesuchte Wahrscheinlichkeit:

$P(4 \text{ Richtige}) = \frac{\binom{6}{4} \cdot \binom{43}{2}}{\binom{49}{6}} \approx 0{,}1\ \%$



c) Analog gilt: $P(0 \text{ Richtige}) = \frac{\binom{6}{0} \cdot \binom{43}{6}}{\binom{49}{6}} = \frac{\binom{43}{6}}{\binom{49}{6}} \approx 0{,}4360$