Zweiseitiger Signifikanztest / Hypothesentest

Einleitung:

Bisher war die Wahrscheinlichkeit p eines Bernoulli-Versuchs bekannt. Eine Möglichkeit p herauszufinden, ist es, einen Versuch so oft wie möglich hintereinander auszuführen. Nach dem Gesetz der großen Zahlen nähert sich dann die relative Häufigkeit an p an. Möchte man also wissen, ob ein Marmeladenbrot wirklich immer auf die Marmeladenseite fällt, so muss man das Brot nur sehr oft auf den Boden fallen lassen. Allerdings kann man die Vermutung auch stichprobenartig überprüfen. Hierzu ist die Binomialverteilung sehr nützlich.

Beispiel: Tom und Peter entscheiden in ihrem Alltag als Freunde vieles per Münzwurf. Tom hat das Gefühl, dass er meistens verliert, wenn z.B. per Münzwurf entschieden wird, wer das nächste Getränk holen muss. Tom, der immer Kopf wählt, vermutet, dass die Münze von Peter gezinkt ist. Dies würde er gerne überprüfen.

Wie kann man nun feststellen, ob Tom Recht hat?

Man wirft die Münze z.B. 20-mal. Wenn die Münze nicht gezinkt ist, erscheint „Kopf" mit einer Wahrscheinlichkeit von p = 0,5, also 10-mal. ($20 \cdot 0{,}5 = 10$). Was ist aber, wenn bei 20 Würfen 9-mal oder 11-mal Kopf erscheint. Dies liegt wohl noch im Toleranzbereich und man würde vermuten, dass Peters Münze nicht gezinkt ist. Was ist aber bei 8mal bzw. 12mal und was bei 7 und 13-mal. Wer hat dann Recht?

Man muss also eine Trefferanzahl bzw. einen Bereich festlegen, in dem man Peter Recht geben kann.

Wie würdest du gefühlsmäßig den Annahmebereich festlegen? Sind die Grenzen gegenüber Peter fair gewählt?

Aufgabe: Erarbeitet in Gruppen ein Beispiel für einen zweiseitigen Hypothesentest. Stellt dazu die Binomialverteilung auch graphisch dar. Benutzt einen Gegenstand aus dem Wahrscheinlichkeitskoffer.

Weiter im Beispiel:

Tom und Peter entscheiden sich für folgende Vorgaben:

Wie vermutet bleibt Peter bei p = 0,5 (Nicht gezinkt).

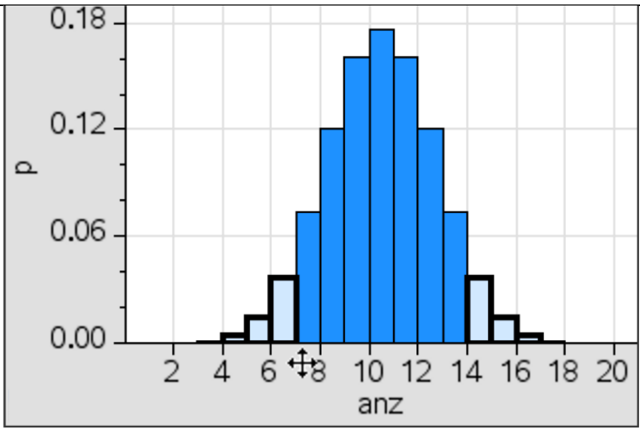
Tom behauptet $p \neq 0{,}5$.

Weiter soll gelten n = 20

Nun wird der Annahmebereich festgelegt.

Als Erwartungswert gilt: $E(x) = \mu = n \cdot p = 20 \cdot 0{,}5 = 10$

Als Annahmeintervall wählen sie I : [8;12] und damit als Ablehnungsbereich linksseitig I_1: [0 ; 7] und rechtsseitig I_2 : [13 ; 20]



Nun wird geworfen: Tom wirft 20-mal und es erscheint 9-mal Kopf. Daraufhin entschuldigt er sich bei Peter und holt zur Versöhnung zwei Bier.

„Glück gehabt!", denkt sich Peter und lässt die gezinkte Münze lieber mal schnell verschwinden.

Sollte also bei der Durchführung die Trefferanzahl im Annahmebereich liegen, bedeutet das nicht, dass Tom dann Unrecht hat, sondern nur, dass seine Vermutung sehr zweifelhaft ist.

Bisher haben wir den Annahmebereich nach Gefühl festgelegt. Es ist klar, dass Mathematiker, dafür lieber eine Formel verwenden möchten.
Nach den Sigma-Regeln gilt, dass mit einer Wahrscheinlichkeit von 95% eine Stichprobe einen Wert im Intervall **[μ - 1,96·σ ; μ + 1,96·σ]** ergibt. Genaueres hierzu steht auf dem Arbeitsblatt: Veranschaulichung der Sigmaregel. Wir bestimmen die Grenzen mit Hilfe der Sigmaregel nach folgendem „Verfahren“:

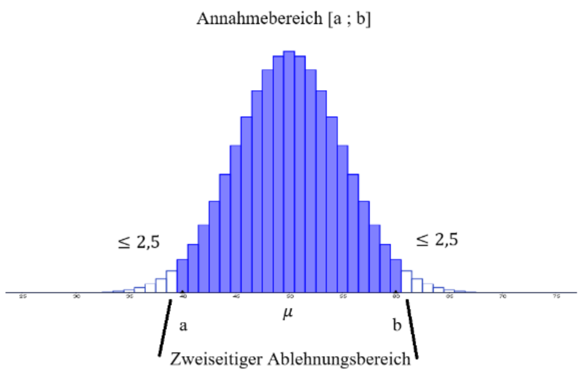
n = 20 E(x) = μ = 20 · 0,5 = 10 $\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{20 \cdot 0{,}5 \cdot 0{,}5} = \sqrt{5} = 2{,}24$

Annahmebereich: [10 - 1,96 · 2,24 ; 10 + 1,96· 2,24] = [10 – 4,38 ; 10 + 4,38] = [5,62 ; 14,38]. Nun runden wir so, dass der Annahmebereich und nicht der Ablehnungsbereich vergrößert wird. Es gilt dann für den Annahmebereich: I = [5 ; 15]. Nicht irritieren lassen, wenn in Musteraufgaben die Grenzen leicht abweichen. Dann wurde meist anders gerundet. Unsere Regel: “Untere Grenze nach unten und obere Grenzen nach oben runden.“

Wir haben nun einen Annahmebereich von mindestes 95% und einen Ablehnungsbereich von höchstens 5% gewählt. Klar, dass Peter, dem wir nur beste Absichten unterstellen, tendenziell eher Recht bekommt. Es ist nicht sinnvoll, bei jeder neuen These die alte gleich zu verwerfen. Diese 5% bezeichnet man auch als Signifikanzniveau. Bei einem Signifikanzniveau von 5% multipliziert man mit 1,96. Man kann sich aber auch für ein anderes Signifikanzniveau entscheiden. Für ein Signifikanzniveau von 1% würden wir beim zweiseitigen Test dann mit 2,58 und für 10% mit 1,64 multiplizieren. Das Signifikanzniveau muss man also vorher festlegen. Klar, wer die Ursprungshypothese (Nullhypothese H_0) vertritt, möchte einen großen Annahmebereich und einen kleinen Ablehnungsbereich und wählt ein geringes Signifikanzniveau. Wer die neue Hypothese (Forschungshypothese H_1) vertritt, hätte gerne einen großen Ablehnungsbereich und einen geringen Annahmebereich und hätte gerne ein großes Signifikanzniveau. Ein Annahmebereich von 90% oder mehr erscheint recht groß. Aber auch in der medizinischen Forschung führt man ein neues Medikament erst ein, wenn es eine deutlich bessere Wirkung erreicht. Daher ist ein recht geringer Ablehnungsbereich sinnvoll.

Eine andere Methode ist die folgende:

Wir legen fest, dass die Wahrscheinlichkeit für den Ablehnungsbereich höchstens 5 % betragen soll. Für den links- und rechtsseitigen Ablehnungsbereich ergibt sich dann höchstens eine Wahrscheinlichkeit von 2,5. Nun sucht man die Zahlen k = a (linksseitig) und b (rechtsseitig), welche als Grenzen noch im Annahmebereich liegen.



Wir suchen also a und b so, dass gilt P(a ≤ X ≤ b) ≥ 0,95 bzw. 95%.

Linksseitige Grenze a	Rechtsseitige Grenze b
Wir wollen höchstens 2,5% vom Annahmebereich abschneiden bzw. der Ablehnungsbereich soll ≤ 2,5% sein. Wir betrachten die kumulierten Tabellen und suchen die größte Zahl bei der 2,5% bzw. 0,025 noch nicht überschritten wird. Das ist dann unsere unterste Grenze, die noch zum Annahmebereich gehört.	Für b rechnen wir kurz: 100% - 2,5% = 0,975 Daher gilt für b: P(X ≤ b) > 0,975. Auch hier betrachten wir wieder die kumulierten Tabellen und suchen die kleinste Zahl b, bei der 0,975 bzw. 97,5% gerade überschritten wird.

Binomialverteilung $F_{n;\,p}(k) = \sum_{i=0}^{k} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}$

Summenverteilung

n	k	0,02	0,03	0,05	0,10	1/6	0,20	0,25	0,30	1/3	0,40	0,50	
20	0	0,6676	5438	3585	1216	0261	0115	0032	0008	0003	0000	0000	19
	1	9401	8802	7358	3917	1304	0692	0243	0076	0033	0005	0000	18
	2	9929	9790	9245	6769	3287	2061	0913	0355	0176	0036	0002	17
	3	9994	9973	9841	8670	5665	4114	2252	1071	0604	0160	0013	16
	4		9997	9974	9568	7687	6296	4148	2375	1515	0510	0059	15
	5			9997	9887	8982	8042	6172	4164	2972	1256	0207	14
	6				9976	9629	9133	7858	6080	4793	2500	0577	13
	7				9996	9887	9679	8982	7723	6615	4159	1316	12
	8				9999	9972	9900	9591	8867	8095	5956	2517	11
	9					9994	9974	9861	9520	9081	7553	4119	10
	10					9999	9994	9961	9829	9624	8725	5881	9
	11						9999	9991	9949	9870	9435	7483	8
	12							9998	9987	9963	9790	8684	7
	13								9997	9991	9935	9423	6
	14									9998	9984	9793	5
	15										9997	9941	4
	16											9987	3
	17	Nicht aufgeführte Werte sind (auf 4 Dez.) 1,0000.										9998	2

Untere Grenze:

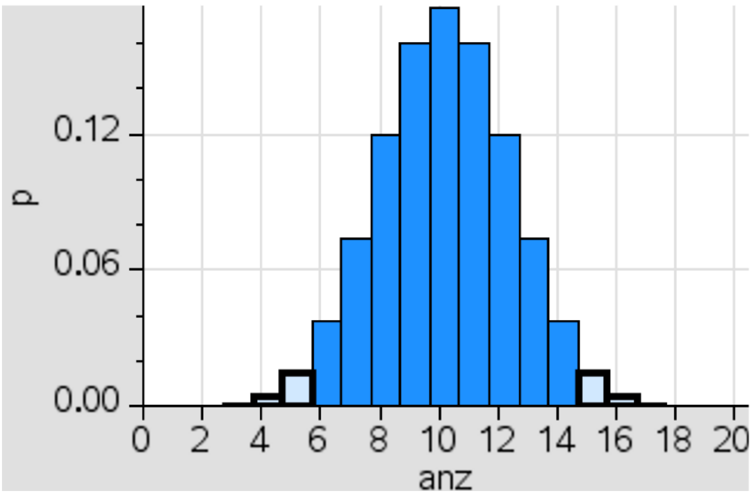
Wir sehen, dass die untere Grenze bei p = 0,5 für den Wert k = 5 gerade noch unterhalb von 2,5% liegt (0,0207 < 0,0205).

Obere Grenze:

Wir sehen, dass die obere Grenze bei p = 0,5 für den Wert k = 14 gerade oberhalb von 95,5% liegt (0,9793 > 0,95).

Damit haben wir die Grenzen für den Annahmebereich: I = [5 ; 14] gefunden. Im Vergleich zum obigen Annahmebereich haben wir eine etwas andere obere Grenze, da wir die Grenzen anders ermittelt haben und auch andere Rundungswerte benutzen.

Damit gilt für den Annahmebereich das Intervall [5 ; 14]



Aufgabe:

Nehmt euer Beispiel aus der Gruppenarbeit und berechnet nun den Annahmebereich, sowohl mit einem Signifikanzniveau von 5% als auch mit den Tabellen.

Zusammenfassung der Begriffe:

Hypothesentest: Beim Hypothesentest überprüft die Richtigkeit einer vermuteten Wahrscheinlichkeit p.

Nullhypothese N_0: Die Ursprungshypothese bezeichnet man als Nullhypothese. Hier p = 0,5.

Alternativhypothese H_1: Die Forschungshypothese bezeichnet man als Alternativhypothese. Hier p ≠ 0,5.

Stichprobe(-numfang): Die Anzahl der Testpersonen bezeichnet man also Stichprobenumfang.

Signifikanzniveau: Das Signifikanzniveau wird vor dem Test festgelegt. Üblicherweise höchstens 1%, 5% oder 10%.

Annahmebereich: Intervall für die Trefferzahl, in dem man die Nullhypothese annimmt.

Ablehnungsbereich: Intervall für die Trefferzahl, in dem man die Nullhypothese verwirft.

Vorgehen beim Entwerfen eines Tests:

1.) Bestimme die Nullhypothese N_0: z.B.: p = 0,5. (Damit auch die Alternativhypothese: Hier p ≠ 0,5.)

2.) Lege eine Stichprobenprobengröße fest. n = 20

3.) Bestimme die Grenzen nach folgender Methode (Je nach Aufgabenstellung)

   a) „Gefühlsmäßig“
   b) Nach den Sigmaregeln: Erwartungswert bestimmen, Sigma berechnen, Bereich berechnen, Grenzen runden.
   c) Nach Tabelle oder GTR: Tabelle öffnen. 1. Folge erzeugen. 2. Binomcdf(n, p), Grenzen suchen

Entwirf einen Hypothesentest mit n = 20 und p = 0,4. Verwende sowohl die Sigmaregel als auch die kumulierten Tabellen. Das Signifikanzniveau soll bei 5% liegen.

1. Nullhypothese H_0 : p = 0,4 (Alternativhypothese: H_1 : p ≠ 0,4

2. Stichprobengröße: n = 20

3. Erwartungswert: $E(x) = \mu = 20 \cdot 0{,}4 = 8$

4. Sigma: $\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{20 \cdot 0{,}4 \cdot (1-0{,}4)} = \sqrt{4{,}8} = 2{,}19$

5. Annahmebereich: **[μ - 1,96·σ ; μ + 1,96·σ] =** $[8 - 1{,}96 \cdot 2{,}19\ ;\ 8+ 1{,}96 \cdot 2{,}19] = [3{,}71;12{,}29] \approx [3;\ 13]$

2 Laura behauptet, dass Lukas mit einem gezinkten Würfel würfelt, der nicht die zu erwartende Anzahl Sechsen würfelt. Um die Behauptung zu testen, wirft sie Lukas' Würfel n-mal.
Wie ist beim Signifikanzniveau 5 % zu entscheiden, wenn dabei k Sechsen fallen?
a) $n = 25$; $k = 6$ b) $n = 50$; $k = 12$ c) $n = 100$; $k = 24$

3 Bei einer Lotterie zieht eine „Lotto-Fee" aus der Urne in Fig. 1 eine Kugel. Falls eine rote Kugel gezogen wird, gewinnt man einen Preis. Ein Spieler zweifelt, ob die Kugel von der Fee wirklich zufällig gezogen wird.
Bestimmen Sie für einen Signifikanztest auf dem Signifikanzniveau 5 % bei einem Stichprobenumfang von $n = 50$ $(n = 100)$ den Annahmebereich für die Hypothese „Die Fee arbeitet einwandfrei". Wie groß ist die Irrtumswahrscheinlichkeit?

4 Eine Partei hatte bei der letzten Wahl einen Stimmenanteil von 40 %. Ein Parteisekretär untersucht, ob sich der Stimmenanteil in seinem Bezirk verändert hat. Er führt einen Signifikanztest auf dem Signifikanzniveau 5 % mithilfe einer repräsentativen Umfrage bei 100 Wählern durch. Davon geben 33 an, dass sie die Partei bei der nächsten Wahl wählen wollen. Welches Ergebnis liefert der Signifikanztest für die Nullhypothese „Der Stimmenanteil ist gleich geblieben"? Beurteilen Sie das Ergebnis.

5 Eine Nussmischung soll 30 % Walnüsse und 70 % Haselnüsse enthalten. Eine Maschine füllt die Nüsse in Tüten von je 50 Nüssen ab. Man greift zwei Tüten heraus und zählt 80 Haselnüsse. Entscheiden Sie mithilfe eines Signifikanztests auf einem Signifikanzniveau von 5 %, ob man diese Abweichung tolerieren kann.

Lösung Grenzen können leicht abweichen.

2 Nullhypothese ist jeweils: H_0: $p = \frac{1}{6}$, in Worten: Lukas hat keinen gezinkten Würfel (nur darauf kann man testen, sonst müsste man die Wahrscheinlichkeit für den gezinkten Würfel kennen).
Die Testvariable X zählt die Sechsen, Parameter sind n (je nach Teilaufgabe) und $p = \frac{1}{6}$.
a) Annahmebereich = [1; 8]. Da k im Annahmebereich liegt, wird die Nullhypothese beibehalten.
b) Annahmebereich = [4; 14]. Da k im Annahmebereich liegt, wird die Nullhypothese beibehalten.
c) Annahmebereich = [10; 24]. Da k im Annahmebereich liegt, wird die Nullhypothese beibehalten.

3 Nullhypothese: Eine rote Kugel wird zufällig gezogen; H_0: $p = 0{,}2$.
Testvariable X: Anzahl der roten gezogenen Kugeln, Parameter $n = 50$ bzw. 100 und $p = 0{,}2$.
Annahmebereich [5; 16] bzw. [12; 28].
Der Annahmebereich ist bei $n = 50$ relativ zu n groß, bei $n = 100$ ist er etwas kleiner.
Da

$$1 - P(5 \le X \le 16) = 1 - P(X \le 16) + P(X \le 4) = 0{,}0329$$

bzw.

$$1 - P(12 \le X \le 28) = 1 - (P(X \le 28) + P(X \le 11) = 0{,}0326$$

sind, beträgt die Irrtumswahrscheinlichkeit jeweils etwa 3,3 %.

4 Nullhypothese: Der Stimmanteil ist gleich geblieben, H_0: $p = 0{,}40$.
Testvariable X: Anzahl der Wähler der Partei mit $n = 100$, $p = 0{,}40$.
Annahmebereich: [31; 50].
Bei dem Stichprobenergebnis 33 wird man aufgrund des Signifikanztests die Nullhypothese nicht verwerfen. Demnach hat sich der Stimmenanteil nicht signifikant verändert.

5 Nullhypothese: Die Mischung enthält 70 % Haselnüsse (und 30 % Walnüsse); H_0: $p = 0{,}7$.
Testvariable X: Anzahl der Haselnüsse mit $n = 100$, $p = 0{,}7$.
80 liegt nicht im Annahmebereich [61; 79], also kann die Abweichung nicht toleriert werden.