1. Drei Grundfragen

In der schliessenden Statistik geht es um die Frage, wie und mit welcher „Sicherheit" von Ergebnissen, die nur an einem Teil der Gesamtheit (Stichprobe) gewonnen wurden, auf die Situation der Grundgesamtheit geschlossen werden kann.

Grundfragen	Schätzungen	Welche Kennzahl, Masszahl (Parameter) passt am besten zu den Beobachtungen?
	Statistische Tests	Sind die Beobachtungen mit einer vorgegebenen Kennzahl (Parameter) vereinbar? Im Beispiel der Abstimmung, könnte der Ja-Anteil auch kleiner als 50% sein? Ist im Beispiel der Asbestfasern der Grenzwert eingehalten?
	Vertrauens-intervalle	Welche Verschiedenen Kennzahlen (Parameter) sind für die gemachten Beobachtungen insgesamt plausibel? Das heisst, wie genau sind die berechneten Kennzahlen? Mit welchen Schätzfehlern ist zu rechnen? Im Beispiel der Abstimmung können Ja-Resultate zwischen 48% und 55% erwartet werden.
	Schliessende Statistik	Es geht um zufällig gemessene Daten und gezogenen Schlüssen aus diesen Daten. Wahrscheinlichkeitsrechnung als Grundlage der schliessenden Statistik: Schätzungen: Für welchen Parameter der Grundgesamtheit sind die gemachten Beobachtungen am wahrscheinlichsten? Tests: Sind die Beobachtungen für einen vorgegebenen Parameter wahrscheinlich oder unwahrscheinlich? Vertrauensintervalle: Für welche Parameter (Intervall) sind die Beobachtungen genug wahrscheinlich (plausibel), für welche eher unwahrscheinlich?

2. Notation, Variabeln

Variabeln-bezeichnung	G	Grundgesamtheit	μ	arithmetisches Mittel von M. Kommen mehrere Merkmale vor x,y,z etc. vor, so schreiben wir μx, μy, μz.
	M	Merkmal (Ist ein Untersuchungsgegenstand → Das Merkmal M besitzt in G eine unbekannte, zu erforschende Häufigkeitsverteilung.)	σ^2	Varianz von M. oder bei mehreren Merkmalen σx^2, σy^2, σz^2
	σ	Standardabweichung von M. Oder bei mehreren Merkmalen σ_x, σ_y, σ_z	π	Im Falle eines nominalen Merkmals entspricht das arithmetische Mittel von M einem Anteilswert, bzw. einer Proportion. In dieser Situation verwenden wir für diese Masszahl den Buchstaben π anstatt μ. Oder bei mehreren Merkmalen π_x, π_y, π_z.
	N	Grösse der Grundgesamtheit, manchmal bekannt, manchmal unbekannt.	$\bar{X}$	arithmetisches Mittel der Daten
Verwendete Veriabeln-bezeichnungen in der Stich-probe S	$x_1, x_2, \ldots, x_n$	gemessene Werte (Daten) des Merkmals M in der Stichprobe	$(sx)^2$	empirische Varianz der Daten, bei mehreren Merkmalen s_x^2, s_y^2, s_z^2
	n	Stichprobengrösse	Sx $\left(\sqrt{(sx)^2}\right)$	empirische Standardabweichung der Daten, bei mehreren Merkmalen s_x, s_y, s_z
	p	Im Falle eines nominalen Merkmals M entspricht x einem Anteilswert, bzw. einer Proportion. In dieser Situation verwenden wir für diese Masszahl den Buchstaben p anstatt x. Manchmal wird auch die Bezeichnung ˆp verwendet. Bei mehreren Merkmalen px ,py ,pz		

Für die Zufallsvariablen verwenden wir Grossbuchstaben, für die Daten kleine Buchstaben.	$X_1, X_2, X_3, \ldots, X_n$	n Zufallsvariablen für die n zufälligen Messungen in der Stichprobe	$E[X_1]$	Erwartungswert der Zufallsvariablen X_1
	$var(X_1)$	Varianz der Zufallsvariablen X_1. Die Wurzel der Varianz ist die Standardabweichung von X_1.	X_1	X_1 kann alle Werte des Merkmals M der Grundgesamtheit G annehmen. Seltene Werte von M haben eine kleinere Wahrscheinlichkeit, häufigere Werte von M eine grössere. **Die Wahrscheinlichkeitsverteilung von X_1 entspricht genau der unbekannten Häufigkeitsverteilung des Merkmals M in G!** Deshalb gilt : $E[X_1] = \mu$ sowie $var(X_1) = \sigma^2$

3. Rechenregeln für Erwartungswert und Varianz

Rechenregeln	1.	$E[X \pm Y] = E[X] \pm E[Y]$	2.	$E[a \cdot X] = a \cdot E[X] \quad a \in \mathbb{R}$	3.	$var(X) = E[X^2] - (E[X])^2$
	4.	$var(a \cdot X) = a^2 \cdot var(X)$	5.	Sind X und Y zwei unabhängige Zufallsvariabeln, das heisst, dass die Wahrscheinlichkeit der Ereignisse $X=x_i$ und $Y=y_i$ sich nicht gegenseitig beeinflussen, so gilt für die Varianz der Summe die einfache Regel (da $var(X+Y) = var(X) + var(Y)$		

4. Der zentrale Grenzwertsatz

In der beschreibenden Statistik würden wir die angegebenen Einkommen $x_1, x_2, \ldots, x_n$ arithmetisch mitteln und den Mittelwert x als das "durchschnittliche Einkommen" ansehen. In der schliessenden Statistik müssen wir die Sache anders sehen. Das arithmetische Mittel x ist eine Realisation der Zufallsvariablen X!

Satz

Die Zufallsvariablen $X_1, X_2, \ldots, X_n$ seien unabhängig und identisch verteilt (i.i.d.), wobei $E[X_i] = \mu$ und $Var(X_i) = \sigma^2$.

Dann gilt für $S = X_1 + X_2 + \ldots + X_n$ sowie $\bar{X} = \frac{S}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n}$:

$E[S] = n \cdot \mu$ $\qquad E[\bar{X}] = \mu$

$Var(S) = n \cdot \sigma^2$ $\qquad Var(\bar{X}) = \frac{\sigma^2}{n}$

Bsp.

Auf einer grossen Hühnerfarm legen Zehntausende von Hühnern der gleichen Rasse täglich ihre Eier. Die Zufallsvariable X bezeichnet das Gewicht (in Gramm) eines zufällig herausgegriffenen Eis. Eine Stichprobe von 25 Eiern ergab folgende Werte: 55 53 53 49 51 64 58 51 48 59 50 54 62 53 50 62 45 53 55 56 53 41 60 52 48

Schätzen Sie den/die Modell-Parameter.

TR / F4 / 1-Var Stats für die gegebene Liste liefert → x = 53.4 Gramm; sX = 5.4314 Gramm

Die produzierten Eier werden auf der Farm in 6er-Kartons abgepackt (zufällig, ohne jede Grössensortierung). Berechnen Sie ausgehend von den in b) geschätzten Werten den Erwartungswert und die Standardabweichung für die Nettogewichte der Eierkartons.

Die Gewichte $X_1, X_2, \ldots, X_6$ der sechs abgepackten Eier sind normalverteilte Zufallsvariablen mit den obigen (geschätzten) Parametern E(Xi) = x und Var(Xi) = s_{X2}.
Das Nettogewicht $S = X_1 + X_2 + \ldots + X_6$ ist selber wieder eine Zufallsvariable.
Erwartungswert: E(S) = 6×E(Xi) = 6×53.4 = 320.4 Gramm / Varianz: Var(S) = 6×Var(Xi) = 6×5.4314² = 177 / Standardabweichung: √Var(S) = √177 = 13.3 Gramm

Zentraler Grenzwertsatz

Der zentrale Grenzwertsatz erklärt die grundlegende Bedeutung der **Normalverteilung** in der Statistik.

Wenn wir über die zugrunde liegende Verteilung der X_i nichts wissen (der zentrale Grenzwertsatz gilt ja für beliebige Verteilungen), gilt als Regel, dass **n mindestens 30** sein muss. Falls das Merkmal M in der Grundgesamtheit normalverteilt ist, ist der Stichprobenmittelwert X **auch für kleine Stichproben** normalverteilt!

Zentraler Grenzwertsatz:

Sind $X_1, X_2, \ldots, X_n$ unabhängige Zufallsvariable, die alle dieselbe Verteilungsfunktion (und somit dasselbe μ und σ^2) besitzen, so lässt sich die Zufallsvariable

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

mit wachsendem n immer besser durch eine **normalverteilte** Zufallsvariable mit **Erwartungswert** μ und **Varianz** $\frac{\sigma^2}{n}$ approximieren.

5. Schätzungen

Das Ziel der schliessenden Statistik ist es, aus den beobachteten Messwerten die unbekannten Grössen der Grundgesamtheit zu schätzen. z.B. Mittelwert μ, Varianz σ^2, Anteilswert π. Diese unbekannten Grössen sind Parameter der zu Grunde liegenden Wahrscheinlichkeitsverteilungen der zufällig gemessenen Werte.

Schätzwerte	Wenn wir eine Stichprobe ziehen, können wir daraus Masszahlen (Stichprobenmittelwert x, Stichprobenvarianz s^2, relative Häufigkeit p) berechnen, die sich als Schätzwerte für die Parameter μ, σ^2, π verwenden lassen.
Punkt-schätzung	So werden wir etwa in unserem Beispiel von Seite 9 (Nettoeinkommen im Kt. Luzern) den Stichprobenmittelwert x als eine Schätzung für den unbekannten Parameter μ der Grundgesamtheit betrachten. Wir nennen dies eine Punktschätzung, weil nur ein einzelner Wert als Schätzung angegeben wird.
Intervall-schätzung	Wegen der Zufälligkeit der Stichprobenbildung ist aber auch x nur ein "zufälliger" Wert, der von Stichprobe zu Stichprobe verschieden sein wird. Es kann deshalb sinnvoll sein, einen ganzen Bereich, in welchem man den gesuchten Parameter vermutet, als Schätzung anzugeben. Dann spricht man von einer Intervallschätzung oder Konfidenzschätzung.

5.1 Punktschätzungen

Erwartungstreu

Das arithmetische Mittel x ,oder genauer die Beobachtung der Zufallsvariable X, ist ein guter Schätzer für den Parameter μ, weil die Realisationen (d.h. die aus verschiedenen Stichproben gewonnenen Werte) wirklich um μ herum streuen:
Ein solcher Schätzer heisst erwartungstreu: sein Erwartungswert stimmt mit dem zu schätzenden Parameter μ überein. Unter gewissen Voraussetzungen ist dieser Schätzer auch optimal. Dies bedeutet, dass dieser Schätzer eine minimale Streuung hat. Kommen jedoch in den Stichprobendaten Ausreisserwerte vor, so kann der Median der Beobachtungen eine bessere Schätzung für μ darstellen. Für nominale Merkmale, bei denen ein Anteilswert π in der Grundgesamtheit geschätzt werden soll, ist der optimale Schätzer die relative Häufigkeit p des Merkmals in der Stichprobe. Auch dieser Schätzer ist erwartungstreu und optimal.
Bemerkung: Wie Sie bereits von früher wissen, muss als Schätzung für die unbekannte Varianz σ^2 der Grundgesamtheit die empirische Varianz s^2 verwendet werden:

$$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Eine Erklärung für den unerwarteten Faktor (n−1) an Stelle von n werden Sie später erhalten. Auch hier ist der Grund die Erwartungstreue. s^2 ist eben ein erwartungstreuer Schätzer für σ^2.

Beispiel

In den Haushalten im Kanton Luzern wurden in einer statistischen Erhebung die monatlichen Ausgaben für diverse Konsumgüter gemessen. Aus einer zufälligen Auswahl von 3500 Privatadressen (Bruttostichprobe) sind 1200 für die Auswertung brauchbare Fragebögen (Nettostichprobe) eingegangen. Für die Konsumgruppe GR1 ergab die Auswertung der Nettostichprobe folgende Werte. arithmetisches Mittel: 959 (Fr.) Standardabweichung: 414 (Fr.)

Berechnen Sie ein 95% Vertrauensintervall für die durchschnittlichen, monatlichen Ausgaben pro Haushalt für die Konsumgruppe GR1 im Kanton Luzern.

***TR / F7 / 2: TInterval** / kennz. / x = 959 / S_x = 414 / n = 1200 / C level = 0.95 → C Int (935.6, 982.4)*

HH Grösse	Stichprobengrösse	ar. Mittel	Standardabweichung	Anteil in %
1	323	543	130.9	33.4%
2	332	897	236.7	30.4%
3	187	1035	288.7	12.6%
4	221	1267	332.3	14.4%
5+	137	1494	413.4	9.2%

Berechnen Sie mit Hilfe der Nachschichtung eine neue, bessere Punktschätzung x_g für die durchschnittlichen, monatlichen Konsumausgaben für GR1 pro Haushalt im Kanton Luzern. (x_g ist ein gewichtetes Mittel, z.B. ist x_1 = 543 repräsentativ für 33.4% aller Haushaltungen.)

xg = 0.334 x 543 + 0.304 x 897 + ... + 0.092 x 1494 = 904.36 ≈ 904

Warum weicht die neue Punktschätzung vom ursprünglichen arithmetischen Mittel von 959 Fr. ab? Unter welchen Umständen würde die Schätzung von b) 959 Fr. ergeben?

Verteilung müsste überall 20% sein. Hier ist sie allerdings gewichtet.

Aus der Simulation einer geschichteten Stichprobe im Kurs der quantitativen Sozialforschung wissen wir, dass eine geeignete Schichtung genauere Schätzungen liefert. Verifizieren Sie diesen Sachverhalt, in dem Sie ein 95% Vertrauensintervall für Ihre Schätzung in Aufgabe b) berechnen und mit dem Resultat von Aufgabe a) vergleichen. Tipp: Berechnen Sie var(x_g) mit den Rechenregeln für die Varianz.

4. Rechenregel: var (a x X) = a^2 x var (X) → var (x_1) = 130.9^2 / 323 = 53.05 / var (x2) = 236.7^2 / 332 = 168.76…..
0.334^2 x 53.05 + 0.3042 x 168.76 + …… = 49.51 → √49.51) = 7.04
→ 95% Intervall 904.36 + 1.96 x 7.04 = 918 Breite Intervall a) 982 – 936 = 46
904.36 – 1.96 x 7.04 = 891 Breite Intervall d) 918 – 891 = 27 Verkleinerung um 41%

5.2 Vertrauensbereiche für arithm. Mittelwerte

Konfidenz-intervalle

Eine Punktschätzung hat den Nachteil, dass sie keinerlei Hinweise darüber enthält, wie "genau" die Schätzung wirklich ist. Die Abweichung zwischen der Punktschätzung und dem wahren Parameter (also z.B. | x −μ |) kann recht gross sein, besonders dann, wenn der Umfang der Stichprobe klein ist. Diesen Mangel behebt die Methode der Konfidenzintervalle (oder Vertrauensbereiche).

Konfidenz-niveau

Ausgehend von einem im Voraus festzusetzenden Fehlerrisiko α wird dabei ein Zufallsintervall [a,b] bestimmt, das den unbekannten Parameter (z.B. den Erwartungswert μ) mit der Wahrscheinlichkeit 1−α überdeckt. Die Wahrscheinlichkeit 1−α heisst dabei das Konfidenzniveau. (**a=5% → 95% Sicherheit; a=1% → 99% Sicherheit**)

Grösse der Stichproben

Grosse Stichprobe	**n > 5% von N** Standartfehler wird $\frac{\sigma}{\sqrt{n}}$ durch $\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ ersetzt. Deshalb nicht mit TR möglich!	Kleine Stichprobe	**n < 5% von N** Die Grösse der Grundgesamtheit hat keinen Einfluss auf den Vertrauensbereich.

Endlichkeits-faktor

Der hier verwendete Korrekturfaktor $\cdot\sqrt{\frac{N-n}{N-1}}$ heisst Endlichkeitskorrektrur. Er kommt daher, dass in einer Stichprobe die Stichprobenelemente ohne Zurücklegen aus der Grundgesamtheit gezogen werden. Es ist zu beachten, dass bei einer Vollerhebung (n=N) der Faktor die Grösse Null annimmt. Bei einer Vollerhebung ist der Standartfehler gleich Null.

Beispiel

Von 100 hergestellten Zementröhren misst man bei 16 zufällig ausgewählten Röhren den Innendurchmesser und stellt ein Stichprobenmittel x = 21.63 cm fest; die Standard-abweichung der Stichprobe beträgt s = 0.3 cm. Welches Intervall enthält mit 99%iger Sicherheit den durchschnittlichen Innendurchmesser aller 100 Röhren?

Weil $\frac{n}{N} = \frac{16}{100} = 0.16 > 0.05$ ist, handelt es sich um eine "grosse" Stichprobe.

Wir berücksichtigen deshalb den Endlichkeitsfaktor.

$$\text{Standardfehler} = \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}} = \frac{0.3}{\sqrt{16}} \cdot \sqrt{\frac{100-16}{100-1}} = \frac{0.3}{4} \cdot \sqrt{\frac{84}{99}} = 0.069085$$

Ablesen der Quantils für die t-Verteilung
($v = 16-1 = 15$, Konfidenzniveau 99%, zweiseitig): $t = 2.947$

$$\bar{x} \pm t \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}} = 21.63 \pm 2.947 \cdot 0.069085 = 21.63 \pm 0.2036$$

Mit 99%iger Sicherheit enthält das Intervall von 21.42 cm bis 21.84 cm den durchschnittlichen Innendurchmesser aller 100 Röhren.

Anzahl Freiheitsgrade

Für grosse Werte von n geht dagegen die t-Verteilung in die Normalverteilung über. Die **Dichtefunktionen** der t-Verteilung (die auch in anderem Zusammenhang Anwendung finden) werden nicht mit Hilfe des Stichprobenumfangs n "nummeriert", sondern in Abhängigkeit der Anzahl Freiheitsgrade v (Nü). Allerdings besteht ein einfacher Zusammenhang: **v = n −1**.

TR

ZInterval	TInterval	1-PropZInt
Stats/Tests → 7: ZInterval Berechnet ein Vertrauensintervall für den Erwartungswert μ der Grundgesamtheit (σ **bekannt**)	Stats/Tests → 8: TInterval Berechnet ein Vertrauensintervall für den Erwartungswert μ der Grundgesamtheit (σ **unbekannt**)	Stats/Tests → A: 1-PropZInt Berechnet ein **Vertrauensintervall** für den Anteilswert π in der Grundgesamtheit
	Der Kanton Luzern wies 2000 ca. 141'000 Haushaltungen auf. Ein Marktforschungsinstitut möchte die durchschnittlichen monatlichen Ausgaben für Süssigkeiten ermitteln und befragt 500 Haushaltungen. Die Stichprobe ergibt x = 135 Fr., s_X = 50 Fr. Welches Intervall überdeckt mit 95.4%iger Sicherheit die durchschnittlichen monatlichen Ausgaben für Süssigkeiten aller Haushaltungen des Kantons Luzern? ***TR / F7 / 2: TInterval** / kennz. / x = 135 / s_X = 50 / n = 500 / C level = 0.954 → C int (130.5, 139.5) → Mit 95%iger Sicherheit überdeckt das Intervall [130.61 Fr., 139.39 Fr.] die durchschnittlichen monatlichen Ausgaben für Süssigkeiten aller Haushaltungen des Kantons Luzern.*	Ein Marktforschungsinstitut befragt auf zufällige Art und Weise 1'200 Haushaltungen in der Schweiz. In 144 Haushaltungen wird das Produkt XY verwendet. Welches Intervall enthält mit 95.5%iger Sicherheit den Anteil p aller Haushaltungen der Schweiz? ***TR / F7 / 5: Z-Int Anteilw** / x = 144 / n = 1'200 / C level = 0.955 → C int (0.1012, 0.1388)* Soll zukünftig weiterhin am Nationalfeiertag eine offizielle Feier auf dem Rütli stattfinden oder nicht? Bestimmen Sie ein 95%-Konfidenzintervall für den Ja-Anteil der ganzen Bevölkerung (D-CH, F-CH). Schreiben Sie einen Antwortsatz. ***TR / F7 / 5: Z-Int Anteilw** / x = 274 / n = 476 / C level = 0.95 → C int (0.5312, 0.62) → Mit 95%iger Sicherheit überdeckt das Intervall [0.531, 0.620] den Ja-Anteil der ganzen Bevölkerung.*

(Zeile „Beispiel“: Die zweite Tabellenzeile enthält die Beispiele.)

5.3 Vertrauensbereiche für Anteilswerte

Mit der zu z gehörenden Wahrscheinlichkeit enthält das Intervall von den Anteil π der Grundgesamtheit.

$$p - z \cdot \sqrt{\frac{p\cdot(1-p)}{n}} \quad \text{bis} \quad p + z \cdot \sqrt{\frac{p\cdot(1-p)}{n}}$$

Die Aussage ist so gültig, wenn **n x p x (1− p) > 9** ist und es sich um eine im Verhältnis zur Grundgesamtheit **kleine Stichprobe** (n/N < 0.05) handelt. Für im Verhältnis zur Grundgesamtheit **grosse Stichproben** (n/N ≥ 0.05) muss dagegen analog zum quantitativen Fall der Korrekturfaktor $\cdot\sqrt{\frac{N-n}{N-1}}$ verwendet werden.

Beispiele

Kleine Stichprobe: (Mit TR möglich) Eine Zufallsstichprobe bei 1'500 erwachsenen Personen der Schweiz ergab 1'080 Personen, die über einen bestimmten politischen Sachverhalt Bescheid wussten. Welches Intervall enthält mit 95%iger Sicherheit den wahren Anteil π, der bei der Befragung aller erwachsenen Personen der Schweiz ermittelt worden wäre?

Kleine Stichprobe

$$p = \frac{1'080}{1'500} = 0.72$$

$$n \cdot p \cdot (1-p) = 1500 \cdot 0.72 \cdot 0.28 = 302.4 > 9 \checkmark$$

$$\text{Standardfehler} \quad \sqrt{\frac{p\cdot(1-p)}{n}} = \sqrt{\frac{0.72\cdot 0.28}{1'500}} = 0.011593$$

Zum Konfidenzniveau $1-\alpha = 0.95$ gehört ein z-Wert von 1.96 (vgl. Seite 16).

$$p \pm z \cdot \sqrt{\frac{p\cdot(1-p)}{n}} = 0.72 \pm 1.96 \cdot 0.011593 = 0.72 \pm 0.0227$$

Mit 95%iger Sicherheit enthält das Intervall von 69.7% bis 74.3% den Anteil aller erwachsenen Personen der Schweiz, welche über den politischen Sachverhalt Bescheid wissen.
Dies ist jedoch mit dem TR möglich und muss nicht von Hand gerechnet werden!!!!
TR / Stats/Tests/ A: 1-PropZInt

Grosse Stichprobe: (nur ohne TR möglich) In einem Schulhaus mit 350 Schülerinnen und Schüler wurden 100 auf zufällige Art und Weise ausgewählt und über die Qualität der Verpflegung in der Kantine befragt. Welches Intervall enthält mit 95%iger Sicherheit den prozentualen Anteil aller 350 Schülerinnen und Schüler, welche die Verpflegung als sehr gut beurteilen?
Sehr gut → 45 Schüler = p = 0.45

$$z \cdot \sqrt{\frac{p\cdot(1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}} = 1.96 \cdot \sqrt{\frac{0.2475}{100}} \cdot \sqrt{\frac{350-100}{350-1}} =$$

= [0.3675, 0,5325]

Mit 95%iger Sicherheit enthält das Intervall von 36.75 bis 53.25% den proz. Anteil aller 350 Schülerinnen und Schüler, welche die Verpflegung als sehr gut beurteilen.

5.4 Stichprobengrösse

Stichproben-grösse

Wenn wir uns eine **maximal zulässige Abweichung F** vorgeben, so stellt sich die Frage, wie gross die Stichprobe mindestens sein muss, damit (bei einem bestimmten Fehlerrisiko α) die gewünschte Genauigkeit erreicht werden kann.

Stichprobengrösse n

$$n \geq \frac{z^2 \cdot \pi \cdot (1-\pi)}{F^2}$$

Beispiele

Im Jahr 1998 wurden bezüglich des Standortes der Messe für Altbausanierung 450 zufällig ausgewählte Besucherinnen und Besucher befragt. 74% der Befragten sprachen sich für eine Beibehaltung des Standortes auf der Allmend Luzern aus. Welche Sicherheit gehört zum Vertrauensbereich 74 +/- 3%?

mit TR Solver nach z aufl.

$$z \cdot \sqrt{\frac{p \cdot (1-p)}{n}} = z \cdot \sqrt{\frac{0.74 \cdot 0.26}{450}} = z \cdot 0.02068 \overset{!}{=} 0.03$$

$z = 1.45$ Tabelle 1 85.3%ige Sicherheit

oder normalcdf(-1.45,1.45) = 0.853

Für das nächste Jahr wird eine erneute Befragung geplant. Wie viele Besucher müsste man mindestens befragen, um bei einer Sicherheit von 95% ein Vertrauensintervall zu erhalten, das nicht breiter ist als 5 Prozentpunkte?

95%ige Sicherheit: $z = 1.96$

maximaler Fehler $F = \frac{\text{Intervallbreite}}{2} = \frac{0.05}{2} = 0.025$

Die Umfrage aus dem Jahr 1998 liefert eine Schätzung für die Standardabweichung:

$\sigma = \sqrt{0.74 \cdot 0.26}$

Somit: minimaler Stichprobenumfang $n = \frac{1.96^2 \cdot 0.74 \cdot 0.26}{0.025^2} = 1'183$ Personen $\qquad n \ge \frac{z^2 \cdot \pi \cdot (1-\pi)}{F^2}$

Oder TR: $Löse(1.96 \cdot \sqrt{\frac{(0.74 \cdot (1-0.74))}{n}} = 0.025, n)$

Der Kanton Luzern wies 2000 ca. 141'000 Haushaltungen auf. Ein Marktforschungsinstitut möchte die durchschnittlichen monatlichen Ausgaben für Süssigkeiten ermitteln und befragt 500 Haushaltungen. Die Stichprobe ergibt x = 135 Fr., s_X = 50 Fr. Wie viele Haushaltungen hätte man befragen müssen, wenn das Vertrauensintervall (bei gleichem Konfidenzniveau) nur eine Breite von 4 Franken hätte betragen dürfen?

95.4%ige Sicherheit → z = 2 *max. Fehler = F = 4 / 2 = 2Standardabweichung = s_X = 50* *minimaler Stichprobenumfang =* $n \ge \frac{z^2 \cdot \pi \cdot (1-\pi)}{F^2}$ $\quad n \ge \frac{2^2 \cdot 50^2}{2^2} = 2500$

Ein Marktforschungsinstitut möchte für einen Kunden den Bekanntheitsgrad seines Markenartikels Z feststellen und fragt, wie viele Personen des mehrere Millionen Konsumenten umfassenden Marktes mindestens befragt werden müssen, damit bei einem Sicherheitsgrad von 95.5% der ermittelte Bekanntheitsgrad um nicht mehr als 3% vom tatsächlichen Bekanntheitsgrad abweicht, wenn…

…keine Schätzung und keine Vorwegstichprobe für den Bekanntheitsgrad existiert.

0.5 = schlechtester Fall!!

$z \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$ 95.5%ige Sicherheit: $z = 2$

$F = 0.03 \quad \sigma = 0.5$ (schlechtester Fall)

$Löse(2 \cdot \sqrt{\frac{0.5 \cdot (1-0.5)}{n}} = 0.03, n)$ oder: $n = \frac{2^2 \cdot 0.5^2}{0.03^2} = 1'111$

… eine Vorwegstichprobe einen Bekanntheitsgrad von 24% ergeben hat.

$z \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$ 95.5%ige Sicherheit: $z = 2$

$F = 0.03 \quad \sigma = \sqrt{0.24 \cdot 0.76}$

$Löse(2 \cdot \sqrt{\frac{0.24 \cdot (1-0.24)}{n}} = 0.03, n)$ oder: $n = \frac{2^2 \cdot 0.24 \cdot 0.76}{0.03^2} = 811$

6. Testen von Hypothesen

Kritischer Bereich K: Auch Verwerfungsbereich. Die Menge aller Ergebnisse, bei deren Eintreffen wir die Nullhypothese verwerfen wollen.

P-Wert: Ist der p-Wert ≤ 0.05 so wird die Nullhypothese verworfen.

Signifikanzniveau: Die Wahrscheinlichkeit P(K) soll klein sein, kleiner als eine Schranke α.
Bei einem Signifikanzniveau von **α = 1%** spricht man von **hochsignifikant**.
Ein Testergebnis, das auf dem Niveau von **α = 5%** zur Ablehnung der Nullhypothese führt, nennt man **signifikant**.

Wenn P-Wert > 0.05 wird H_0 beibehalten
Wenn P-Wert < 0.05 dann wird H_0 verworfen = signifikant
Wenn P-Wert Wert < 0.01 dann wird H_0verworfen = hochsignifikant

Vorgehen beim Testen von Hypothesen

1.	Eine Vermutung kann nicht direkt bewiesen werden → indirekte Bestätigung
2.	Der zu bestätigenden "Alternativhypothese" H_A wird die Nullhypothese H_0 gegenübergestellt, in der Hoffnung, H_0 widerlegen zu können.
3.	Zufallsexperiment planen: Wahl einer geeigneten Testgrösse.
4.	[Evtl. Signifikanzniveau α wählen,] Verwerfungsbereich K bestimmen, Entscheidungsregel aufstellen.
5.	Zufallsexperiment durchführen, Wert der Testgrösse bestimmen. Es sind 2 Entscheidungen möglich: a) Wert der Testgrösse fällt in den Verwerfungsbereich: H_0 wird verworfen. (**Fehlerrisiko 1. Art**: H_0 wird verworfen, obwohl H_0 richtig ist. Die WS, diesen Fehler zu begehen, ist kleiner als α.) b) Wert der Testgrösse fällt nicht in den Verwerfungsbereich: H_0 wird beibehalten. (**Fehlerrisiko 2. Art**: H_0 wird beibehalten, obwohl H_0 falsch ist. Die WS β für diesen Fehler kann im Allgemeinen nur durch zusätzliche Überlegungen oder Annahmen berechnet werden.

Beispiel

Eine Wahrsagerin behauptet von sich selber, übersinnliche Fähigkeiten zu besitzen: sie könne mit einer Treffsicherheit von mindestens 75% das Ergebnis eines Münzenwurfs im Voraus prophezeien. Die Behauptung der Wahrsagerin soll mit einem Test überprüft werden. Sie soll das Ergebnis von 40 Münzenwürfen jeweils im Voraus angeben.
Wie ist der kritische Bereich K zu wählen, damit das Fehlerrisiko 1. Art kleiner als 5% ausfällt?

Nullhypothese H_0: Anzahl X der richtig vorausgesagten Münzwürfe ist binomial verteilt mit Parameter p = 0.75
Der kritische Bereich K = {0,1, 2,..,k} ist so zu wählen, dass bei Gültigkeit von H_0 die Werte von X nur mit einer Wahrscheinlichkeit von maximal 5% in K zu liegen kommen.
Wähle also das grösstmögliche k mit der Eigenschaft, dass P(X ≤ k) < 0.05. Z.B. mit Ausprobieren: cat.F3: BinIwkt(40,0.75, 23) = 0.011561
BinIwkt((40,0.75, 24) = 0.026245 < 0.05 → kritischer Bereich K= {0,1,2,...,24}
BinIwkt((40,0.75, 25) = 0.054437 > 0.05

Mit welcher Wahrscheinlichkeit würde eine nur ratende Person dann diesen Test bestehen?
Ratende Person: Anzahl richtiger Voraussagen X ~ B(40,0.5)
Würde diese Person in obiger Testanlage mindestens 25 Münzenwürfe richtig voraussagen, so würde man die Nullhypothese H_0 nicht verwerfen (Fehler 2. Art).
Das geschieht mit der Wahrscheinlichkeit: P(X ≥ 25) =1− P(X ≤ 24) = 1 − binomcdf(40,0.5,24) = 1 − 0.923 = 0.077 = 7.7%

6.3 Tests für Anteilswerte

Einstichprobentest

Frage	Ist in einer Grundgesamtheit der Anteil π für die Ausprägung eines Merkmals verschieden von einem vermuteten Wert π_o?		
Nullhypotese H_0	$\pi = \pi_0$	Alternativhypothese H_A	$\pi \ne \pi_0$ → zweiseitiger Test $\pi > \pi_0$ bzw. $\pi < \pi_0$ → einseitiger Test
Stichprobe	$n > \frac{9}{\pi_0 \cdot (1-\pi_0)}$ $p = \frac{x}{n}$ Anteil des Merkmals in der Stichprobe	Testgrösse	$z = \frac{(p-\pi_0)}{\frac{\sigma}{\sqrt{n}}} = \frac{(p-\pi_0)\cdot\sqrt{n}}{\sigma}$ mit $\sigma = \sqrt{\pi_0 \cdot (1-\pi_0)}$ ist angenähert **standardnormalverteilt**
TR	F6 / 5: 1-PropZTest		

Beispiel

Gemäss Aussage von Konzernleiter Ulrich Gygi am 7.9.2004 ist das Ziel der Post, die Wartezeiten an den Postschaltern zu verkürzen, erreicht worden: in 92% der Fälle werde der Kunde innerhalb von 7 Minuten bedient. Eine Stichprobe zu verschiedenen Zeiten an zufällig ausgewählten Schaltern ergibt: von 200 Kunden mussten 22 länger als 7 Minuten anstehen. Kann man angesichts dieses Resultats die Aussage des Postchefs glauben?
Nullhypothese H_0: Der Anteil π der Kunden, die am Schalter länger als 7 Minuten warten müssen, beträgt 8%. π = 0.08
Alternativhypothese H_A: Der Anteil p der Kunden, die am Schalter länger als 7 Minuten warten müssen, ist grösser als 8%. π > 0.08 (einseitiger Test)
TR / F6 / 5: Z-TestAnteilw.. p_o = 0.08 / x = 22 / n = 200 / prop > p_o → z = 1.56386; P Value = 0.058925 → 0.058925 > 0.05 H_0 kann nicht verworfen werden. Die Aussage des Postchefs wird durch das beobachtete Ergebnis statistisch nicht widerlegt.

Zustimmungsantiel Armeewaffen – Frauenorganisationen behaupten, dass mehr als $^2/_3$ der Frauen diese Initiative annehmen würden.
Nullhypothese H_0: Der Zustimmungsanteil π der Frauen beträgt $^2/_3$. π = $^2/_3$
Alternative H_A: Der Zustimmungsanteil π der Frauen ist grösser als $^2/_3$. π > $^2/_3$
TR / F6 / 5: Z-TestAnteilw.. / p_o = $^2/_3$ / x = (0.755x388) 293 / n = 388 / prop > p_o → z = 3.7; P Value = 0.000109 → 0.000109 < 0.05 H_0 wird verworfen. Der Zustimmungsanteil der Frauen liegt hochsignifikant höher als $^2/_3$.

Zweistichprobentest

Frage	Sind die Anteile π_1 und π_2 einer Merkmalsausprägung in zwei Grundgesamtheiten verschieden?		
Nullhypotese H_0	$\pi_1 = \pi_2$	Alternativhypothese H_A	$\pi_1 \ne \pi_2$ → zweiseitiger Test $\pi_1 > \pi_2$ bzw. $\pi_1 < \pi_2$ → einseitiger Test
Stichproben	Stichprobenumfänge n_1, n_2 $p_1 = \frac{x_1}{n_1}$ Anteil des Merkmals in der 1. Stichprobe $p_2 = \frac{x_2}{n_2}$ Anteil des Merkmals in der 2. Stichprobe $p = \frac{x_1 + x_2}{n_1 + n_2} = \frac{p_1 \cdot n_1 + p_2 \cdot n_2}{n_1 + n_2}$ Schätzung für π Bedingungen: $(n_1 + n_2) \cdot p \cdot (1-p) > 9 \wedge n_1 \ge 50, n_2 \ge 50$	Testgrösse	$z = \frac{p_1 - p_2}{\sqrt{p \cdot (1-p) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ ist angenähert **standardnormalverteilt**
TR	F6 / 6: 2-PropZTest		

Beispiel

Die Partei ABC erhielt bei einer Umfrage vom Umfang $n_1 = 360$ zum Zeitpunkt t_1 156 Stimmen. Bei einer zum Zeitpunkt t_2 durchgeführten Umfrage hat die Partei 199 von $n_2 = 510$ Stimmen erhalten. Gibt es einen signifikanten Unterschied zwischen den Wähleranteilen zu den beiden Zeitpunkten t_1 und t_2?

Nullhypothese H_0: Die Wähleranteile sind in beiden Zeitpunkten gleich gross. $\pi_1 = \pi_2$
Alternativhypothese H_A: Die Wähleranteile sind in beiden Zeitpunkten unterschiedlich. $\pi_1 \neq \pi_2$ (zweiseitiger Test)
TR / F6 / 6: ZTest2Anteilw.. / $x_1 = 156$ / $n_1 = 360$ / $x_2 = 199$ / $n_2 = 510$ / $p_1 \neq p_2$ → $z = 1.275$; P Value = 0.20288 → 0.20288 > 0.05 H_0 kann nicht verworfen werden. Man wird weiterhin davon ausgehen, dass die Wähleranteile zu den Zeitpunkten 1 und 2 gleich gewesen sind.

Zustimmung Armeewaffen → ist der Zustimmungsanteil für die Intiative in der D-CH und F-CH unterschiedlich?
Nullhypothese H_0: Der Zustimmungsanteil π_1 in der D-CH und π_2 in der W-CH sind gleich gross. $\pi_1 = \pi_2$
Alternative H_A: Die Anteile π_1 und π_2 sind unterschiedlich. $\pi_1 \neq \pi_2$ (zweiseitiger Test)
TR / F6 / 6: ZTest2Anteilw / x_1 = (0.644x527) 339 / $n_1 = 527$ / x_2 = (0.679x202) 137 / $n_2 = 202$ $p_1 \neq p_2$ → $z = -0.89$; P Value = 0.375 → 0.375 > 0.05 H_0 kann nicht verworfen werden. Ein signifikanter Unterschied der Zustimmungsanteile kann nicht nachgewiesen werden.

Gibt es in der Schweiz bzgl. dieser Frage einen Stadt-Land-Unterschied (Stadt: Agglom.=ja, Land: Agglom.=nein)?
Nullhypothese H_0 Der Zustimmungsanteile π_1 in städtischen Gebieten und π_2 in ländlichen Gebieten sind gleich gross. $\pi_1 = \pi_2$
Alternative H_A: Die Anteile π_1 und π_2 sind unterschiedlich. $\pi_1 \neq \pi_2$ (zweiseitiger Test)
TR / F6 / 6: ZTest2Anteilw / x_1 = (0.683x547) 374 / $n_1 = 547$ / x_2 = (0.588x213) 125 / $n_2 = 213$ $p_1 \neq p_2$ → $z = 2.526$; P Value = 0.011537 → 0.011537 < 0.05 H_0 wird verworfen. Der Unterschied der Zustimmungsanteile zwischen Stadt und Land ist signifikant.

Wir betrachten nun die junge Bevölkerungsschicht (15-34). Unterscheidet sich der Zustimmungsanteil dieser Schicht von der restlichen Bevölkerung?
Nullhypothese H_0: Der Zustimmungsanteile π_1 in der Alterschicht 15-34 Jahre und π_2 in der Alterschicht 35-74 Jahre sind gleich gross. $\pi_1 = \pi_2$
Alternative H_A: Die Anteile π_1 und π_2 sind unterschiedlich. $\pi_1 \neq \pi_2$ (zweiseitiger Test)
TR / F6 / 6: ZTest2Anteilw / x_1 = (67% von 208) 139 / $n_1 = 208$ / $x_2 = 360$ / $n_2 = 553$ $p_1 \neq p_2$ → $z = 0.446987$; P Value = 0.654884 → 0.654884 > 0.05 H0 kann nicht verworfen werden. Es ist kein signifikanter Unterschied im Zustimmungsanteil nachweisbar.

Rütli
Der Ja-Anteil ist in der D-CH und F-CH unterschiedlich.
Nullhypothese H_0: Die Ja-Anteile π_1 in der D-CH und π_2 in der W-CH sind gleich gross. $\pi_1 = \pi_2$
Alternative H_A: Die Anteile π_1 und π_2 sind unterschiedlich. $\pi_1 \neq \pi_2$ (zweiseitiger Test)
TR / F6 / 6: ZTest2Anteilw / $x_1 = 199$ / $n_1 = 365$ / $x_2 = 75$ / $n_2 = 111$ / $p_1 \neq p_2$ →
$z = -2.43541$; P Value = 0.014875
→ 0.014875 < 0.05 H_0 wird verworfen. Der Anteil π_1 ist signifikant unterschiedlich vom Anteil π_2

	Total	Region		Wirtschaftsregion			Agglomeration		Alter		
		D-CH 2-4	F-CH 1	A/VA 2	W-ML 3	O-ML 4	ja	nein	15-34	35-54	55-74
TOTAL INTERVIEWS	476	365	111	121	107	137	337	139	155	190	131
1 = ja, es soll weiterhin eine offizielle Feier stattfinden	57.5	54.5	67.6	55.3	51.7	55.9	57.8	56.9	69.7	50.6	53.2
2 = nein, es soll keine Feier mehr stattfinden	35.7	39.3	24.2	41.7	41.5	35.3	34.3	39.1	24.9	39.7	42.8
weiss nicht \ keine Antwort	6.7	6.3	8.2	3.0	6.8	8.8	7.9	3.9	5.4	9.6	4.0
TOTAL ANTWORTEN IN PROZENT	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Mittelwert	1.38	1.42	1.26	1.43	1.45	1.39	1.37	1.41	1.26	1.44	1.45

Ist der Ja-Anteil in der jüngeren Bevölkerungsschicht grösser als in der restlichen Bevölkerung?
Nullhypothese H_0 Die Ja-Anteile π_1 in der Alterschicht 15-34 Jahre und π_2 (35-74) sind gleich gross. $\pi_1 = \pi_2$
Alternative H_A: Der Anteil π_1 ist grösser als π_2. $\pi_1 > \pi_2$ einseitiger Test
TR / F6 / 6: ZTest2Anteilw / $x_1 = 108$ / $n_1 = 155$ / $x_2 = 166$ / $n_2 = 321$ / $p_1 > p_2$ → $z = 3.716$; P Value = 0.000101
→ 0.000101 < 0.05 < 0.01 H0 wird verworfen Der Ja-Anteil der „Jungen" ist hochsignifikant grösser als der Ja-Anteil der „Älteren".

6.4 Chi-Quadrat Unabhängigkeitstest

Um festzustellen ob zwei qualititaive Merkmale vonaneinder unabhängig sind, eignet sich der Chi-Quadrat-Unabhängigkeitstest

Null-hypotese H_0	Unabhängigkeit	Alternativ-hypothese H_A	Abhängigkeit	TR	F6 / 8: Chi2 2-way / Observed Mat → Paste

Beispiele

Bezüglich der Merkmale "Berufsgruppe" und "sportliche Betätigung" wurden 1'000 Personen befragt. Das Ergebnis zeigt die folgende Tabelle. Testen Sie auf einem Signifikanzniveau a = 1%, ob der Grad der sportlichen Betätigung von der Berufsgruppe unabhängig ist.

Berufsgruppe	sportliche Betätigung: nie	gelegentlich	regelmässig
Arbeiter	240	120	70
Angestellter	160	90	90
Beamter	30	30	30
Landwirt	37	7	6
sonstiger freier Beruf	40	32	18

Nullhypothese H_0: Die sportliche Betätigung ist von der Berufsgruppe unabhängig.
Alternativhypothese H_A: Die sportliche Betätigung ist von der Berufsgruppe abhängig.

$$\begin{bmatrix} 240 & 120 & 70 \\ 160 & 90 & 90 \\ 30 & 30 & 30 \\ 37 & 7 & 6 \\ 40 & 32 & 18 \end{bmatrix}$$

TR / F6 / 8: Chi2 2-way / Copy Paste Matrix / $Chi^2 = 38.5541$ / P Value = 0.000006 / df = 8 → 0.000006 < 0.01 → H_0 verwerfen.
Der Grad der sportlichen Betätigung ist hochsignifikant von der Berufsgruppe abhängig

!!!!!! Zur Abklärung der Wirksamkeit einer Grippeimpfung wurden 800 Personen überwacht. Von diesen waren 228 geimpft, 572 nicht. Es wurden folgende Beobachtungen gemacht: Die Frage ist, ob die Impfung gegen Grippe wirklich hilft?

	Erkrankt	nicht erkrankt
geimpft	10	218
nicht geimpft	55	517

Die Kreuztabelle hat die Dimension 2x2. Die benötigte χ^2 -Verteilung für den Unabhängigkeitstest hätte somit nur einen Freiheitsgrad. Ohne Yates-Korrektur wäre dieser Test nicht gültig. Wir weichen deshalb auf den Z-Test für relative Häufigkeiten aus.
Nullhypothese H_0: Der Anteil π_1 der geimpften und erkrankten und der Anteil π_2 der nicht geimpften und erkrankten sind gleich gross: $\pi_1 = \pi_2$
Alternative H_A: Der Anteil π_1 ist kleiner als π_2: $\pi_1 < \pi_2$ (einseitiger Test)
TR / F6 / 6: ZTest2Anteilw / $x_1 = 10$ / $n_1 = 228$ / $x_2 = 55$ / $n_2 = 572$ / $p_1 < p_2$ → $z = -2.4438$; P Value = 0.0073 → 0.0073 < 0.01 H_0 verwerfen. Der Anteil der Geimpften und Erkrankten ist hochsignifikant tiefer als der Anteil der nicht Geimpften und Erkrankten. Die Impfung ist hochsignifikant wirkungsvoll.

Zu Beginn eines Kurses wurden die 180 Teilnehmenden zufällig in drei Gruppen eingeteilt. Die einzelnen Gruppen wurden mit verschiedenen Methoden unterrichtet. Die gemeinsame Abschlussprüfung brachte folgendes Ergebnis: Vor der Prüfung wurde vermutet, dass die verschiedenen Methoden zu unterschiedlichen Lernerfolgen führen würden. Testen Sie diese Vermutung auf einem Signifikanzniveau von 5%.

Gruppe	ungenügend	genügend	gut	sehr gut
1	9	36	9	6
2	12	39	6	3
3	24	33	3	0

→ Die absoluten Häufigkeiten in der letzten Spalte sind zu klein. Deshalb wird die 3. und 4. Spalte zusammengefasst. Dies stellt man fest indem die Matrix in eingegeben wird und dann Exp Mat (15, 36, 6, 3)

$$\begin{bmatrix} 9 & 36 & 15 \\ 12 & 39 & 9 \\ 24 & 33 & 3 \end{bmatrix}$$

TR / F6 / 8: Chi2 2-way / Copy Paste Matrix / $Chi^2 = 16.9$ / P Value = 0.002021 / df = 4 → 0.002021 < 0.05 < 0.01 → H_0 verwerfen.
Lernerfolg und Methode sind hochsignifikant abhängig.