Formelsammlung Stochastik

26. Januar 2013

1. Wahrscheinlichkeit

- Ω: Ereignisraum, Grundraum
- ω: Elementarereignis (sich gegenseitig ausschliessende Ergebnisse)
- A, B, C: Ereignis, Teilmenge von Ω
- Schnittmenge: $A \cap B$ heisst " A und B "
  Falls $A \cap B = \emptyset$: A,B *disjunkt*
- Vereinigung: $A \cup B$ heisst " A oder B ".
- Komplement: $A^c = \overline{A}$ heisst "nicht A"
- Differenz: $A \backslash B = A \cap B^c$ heisst " A ohne B "

1.1. Laplace Modell

uniforme Verteilung:

$P[\omega] = 1/\Omega$ für alle ω

Alle Elementarereignisse haben gleiche Wahrscheinlichkeit!

$$P[A] = \frac{|A|}{|\Omega|} = \frac{\#\omega \in A}{\#\omega \in \Omega}$$

1.2. De Morgan

$$(A \cap B)^c = A^c \cup B^c$$
$$(A \cup B)^c = A^c \cap B^c$$
$$A \cup B = B \cup A$$

1.3. Rechenregeln

$$P[A^c] = 1 - P[A]$$
$$P[A^c] + P[A] = 1$$
$$P[A \cup B] = P[A] + P[B] - P[A \cap B] = P[A \cap B^C] + P[B]$$
$$P[A \cap B] = P[A|B] \cdot P[B] = P[B|A] \cdot P[A]$$
$$P[A \cap B]^c = P[A^c] + P[B^c]$$

1.4. Axiome

$0 \leqslant P[A] \leqslant 1$

$P[\Omega] = 1 \qquad P[\emptyset] = 0$

$P\left(\cup_{i=1}^{m} A_i\right) = P[A_1 \cup ... \cup A_m] \leqslant P[A_1] + ... + P[A_m]$

(wenn ausschliessend)

1.5. Unabhängigkeit

Eintreten von A beeinflusst WK von B nicht. (kein kausaler Zusammenhang.)

$$\boxed{P[A \cap B] = P[A] \cdot P[B]}$$

$P[A \cap B \cap C] = P[A] \cdot P[B] \cdot P[C]$

disjunkt $\Rightarrow$ abhängig $\quad (P[A \cap B] = 0 \neq P[A] \cdot P[B])$

1.6. Kombinatorik

n: # in Grundmenge / k: # in Elemente

$$\boxed{\begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{k!(n-k)!}}$$

	Variation mit Reihenfolge	Kombination ohne Reihenfolge
mit Zurücklegen	n^k	$\binom{n+k-1}{k}$
ohne Zurücklegen	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

Hypergeometrische Verteilung: Total N, betrachte C, davon K richtig in $B \subset N$, $C - K$ richtig in $N - B$ $\qquad \dfrac{\binom{B}{K} \cdot \binom{N-B}{C-K}}{\binom{N}{C}}$

1.7. Zufallsvariable

Funktion von Ω nach $\mathbb{R} : \omega \rightarrow X(\omega)$

- X **diskret**, falls ω „abzählbar"
- X **stetig**, falls ω „in einem Intervall"

W := Wertebereich der Zufallsvariable

2. Gemeinsame und bedingte Wahrscheinlichkeit

2.1. Bedingte Wahrscheinlichkeit

$$\boxed{P[A|B] = P[A \text{ gegeben } B] = \frac{P[A \cap B]}{P[B]}}$$

$$P[A|B] \cdot P[B] = P[A \cap B] = P[B \cap A] = P[B|A] \cdot P[A]$$

Falls A und B unabhängig:

$$P[A|B] = P[A|B^c] = P[A]$$

Rechenregeln:

$P(B|B) = 1$

$P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2)$

$P(A^c|B) = 1 - P(A|B)$

$P(G \cap (G \cup B)) = P(G), \quad P(B \cap (G \cup B)) = P(B)$ Achtung!

$P(A|B) \neq P(B|A) \qquad P(A|B^c) \neq 1 - P(A|B)$

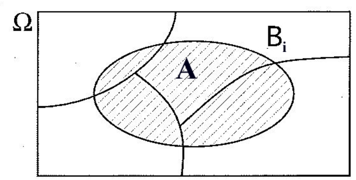
2.2. Satz der totalen Wahrscheinlichkeit

Für k disjunkte Ereignisse $B_1, \ldots, B_k$, wobei $B_1 \cup \ldots \cup B_k = \Omega$ (Partitionierung):

$$\boxed{P(A) = P\left(\bigcup_{i=1}^{k} (A \cap B_i)\right) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P[A|B_i] \cdot P[B_i]}$$

2.3. Satz von Bayes

$$\boxed{P[B_i|A] = \frac{P[A|B_i] \cdot P[B_i]}{P[A|B_1] \cdot P[B_1] + ... + P[A|B_k] \cdot P[B_k]}, \quad B_j \text{ disjunkt}}$$



3. Diskrete Verteilung

3.1. kumultative Verteilungsfunktion

$$F_x(b) = P[X \leqslant b] = P(X \in (-\infty, b]) = \sum_{x_i < x} p(x_i)$$

$P[X \in (a, b]] = F(b) - F(a)$
$P(X > a) = 1 - P(x \in (\infty, a]) = 1 - F(a)$

1. F_x ist **monoton steigend**: $F_x(x_1) \leqslant F_x(x_2)\ \forall\ x_2 > x_1$
2. F_x ist **rechtsstetig**: $\lim_{h\to 0} F_x(x+h) = F_x(x)$
3. $\lim_{x\to\infty} F_x(x) = 1$ und $\lim_{x\to 0} F_x(x) = 0$

3.2. Wahrscheinlichkeitsfunktion p

$p(x_i) = P[X = x_i] = P(\{\omega : X(\omega) = x_i\})$

Eigenschaften: $\quad p(x_i) \geqslant 0, \qquad \sum p(x_i) = 1$

3.3. Erwartungswert und Varianz (diskret)

Kennzahlen für die Verteilung von X.

Erwartungswert (mittlere Lage):

$$\mu_x = E(X) = \sum_{x_i \in W} x_i \cdot p(x_i), \qquad E(g(X)) = \sum_{x_k \in W} g(x_k) \cdot p(x_k)$$

Varianz (Streuung):

$$\boxed{Var(X) = \sigma_x^2 = E[(X - E(X))^2] = \sum (x_i - \mu_i)^2 p(x_i)}$$

$$\boxed{Var(X) = E(X^2) - E^2(X)}$$

Standardabweichung:

$$\boxed{\sigma_x = \sqrt{Var(X)}}$$

3.4. Rechenregeln

$E[a + bX] = a + bE[X] \qquad E[X + Y] = E[X] + E[Y]$

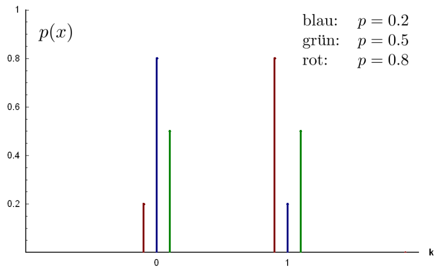
$Var(X) = E[X^2] - E[X]^2$

$Var(a + bX) = b^2 Var(X)$

$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$

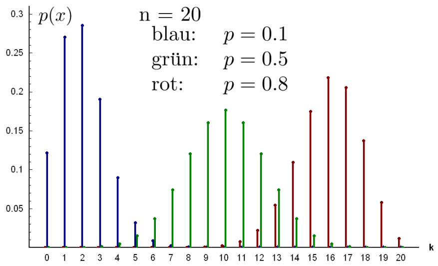
3.5. Bernoulli

Experimente mit ja/nein Ergebnis, $p(x = \{0, 1\})$
$X_k(\omega) =$ Resultat der k-ten Wiederholung
$P(X_k = 1) = p, \quad P(X_k = 0) = 1 - p$



3.6. Binomialverteilung $X \sim Bin(n, p)$

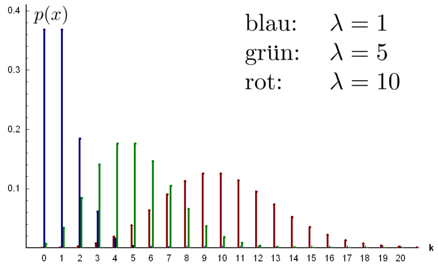
Wiederholung eines Bernoulliexperimentes
#Wiederholungen: n, WK Erfolg: p, #Erfolge: x



3.7. Poissonverteilung $X \sim POI(\lambda)$

n gross, p klein: $\text{Bin}(n, p) \approx \text{Poisson}(\lambda = np)$
Reihe von Ereignissen:
Anzahl pro Zeiteinheit: x, Schnitt pro Zeiteinheit: λ
Bsp: Radioaktiver Zerfall, Callcenter

$X \sim POI(\lambda_1) \quad Y \sim POI(\lambda_2) \quad (X + Y) \sim POI(\lambda_1 + \lambda_2)$

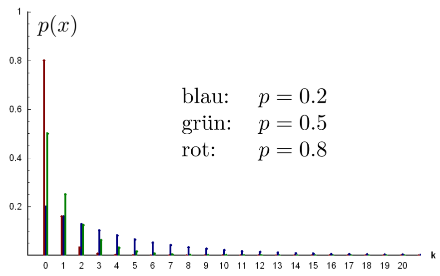


3.8. geometrische Verteilung $X \sim GEO(p)$

$T(\omega) = \min\{k = 1, 2, \ldots : X_k(\omega = 1\}$

$F(x) = \mathbb{P}(X \leq x) = \sum_{i=1}^{x} p(i) = 1 - (1 - p)^x$

Anzahl Wiederholungen x bis Erfolg für Ereignis mit WK p.



3.9. Übersicht diskreter Verteilungsfunktionen

Verteilung	$p(x)$	W_X	$\mathbb{E}[X]$	σ^2 $Var(X)$
Bernoulli(p)	$p^x(1-p)^{1-x}$	{0,1}	p	$p(1-p)$
Bin(n, p)	$\binom{n}{x} p^x (1-p)^{n-x}$	$\{0, \ldots, n\}$	np	$np(1-p)$
Geom (p)	$p(1-p)^{x-1}$	$\{1, 2, \ldots\}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Pois(λ)	$\frac{\lambda^x}{x!} e^{-\lambda}$	$\{0, 1, \ldots\}$	λ	λ

4. Stetige Verteilung

Da $P[X = x] \approx 0$ betrachtet man $P[x \leq X \leq x + h]$

Wahrscheinlichkeitsdichte f:

$$f(b) = F'(b) = lim_{h\to 0} \frac{P[x \leqslant b \leqslant x + h]}{h}$$

Mit der **kumulativen Verteilungsfunktion** $F(x)$**:**

$$F_x(b) = P[X \leqslant b] = \int_{-\infty}^{b} f_x(u) du$$

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x)\, dx$$

Trafo: $F_X^2(x) = P[X^2 \leq x] = P[X \leq \sqrt{x}] = \int_{-\infty}^{\sqrt{x}} f_X(x)\, dx$

Eigenschaften:

1. $f(x) \geqslant 0 \quad F$ ist steigend.
2. $\int_{-\infty}^{\infty} f(x) dx = 1, \quad F(-\infty) = 0 \quad F(\infty) = 1$
3. $f(x) > 1$ ist möglich!

4.1. Erwartungswert und Varianz (stetig)

$$E[X] = \mu_x = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f(x)dx$$

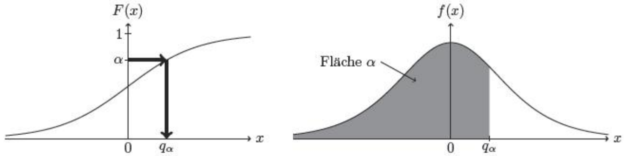
$$Var(X) = \sigma_x^2 = \int_{-\infty}^{\infty} (x-\mu_x)^2 \cdot f(x)dx$$

$$= E\left((x - E[X])^2\right) = E[X^2] - E[X]^2$$

4.2. Quantile $q(\alpha)$

$P[X \leqslant q(\alpha)] = \alpha \qquad q(\alpha) = F^{-1}(\alpha)$

Median wenn $\alpha = 1/2$: Verteilung symmetrisch $\Rightarrow q_{1/2} = E(X)$
Median ist robuster gegen Ausreisser als Erwartungswert



4.3. Verteilungen

4.3.1. Uniforme Verteilung $X \sim UNI(a,b)$

Völlige Ignoranz, benutzt für Messfehler

$W = [a,b]$

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{else} \end{cases} \qquad F[X] = \frac{x-a}{b-a} \quad x \in [a,b]$$

$$E[X] = \frac{(a+b)}{2} \qquad Var(X) = \sigma_X^2 = \frac{(b-a)^2}{12}$$

4.3.2. Exponentialverteilung $X \sim EXP(\lambda)$

Einfachstes Modell für Wartezeiten auf Ausfälle, $W = [0, \infty)$.
Lebensdauer eines Bauteils $T = \frac{1}{\lambda} \rightarrow \mathbb{P}(T \leq 10) = F(10)$
Zeiten zw. Ausfällen eines Systems $\sim Exp(\lambda)$
$\Rightarrow$ Anzahl Ausfälle im Intervall der Länge t $\sim$ $Poisson(\lambda t)$

$$f(x) = \lambda e^{-\lambda x}, x \in (0,\infty) \qquad F[X] = 1 - e^{-\lambda x}\ {}_{[0,\infty]}$$

$$E[x] = \sigma_X = \frac{1}{\lambda} \qquad Var(X) = \sigma^2 = \frac{1}{\lambda^2}$$

Ausfall von erster von n Komponenten $P[x] = 1 - e^{-n\lambda x}$

Summe von Exponentialverteilungen sind *nicht* Exp-verteilt!

Für $X \sim EXP(\lambda_1), \quad Y \sim EXP(\lambda_2)$ unabhängig:

$$F_{X+Y}(c) = 1 + \frac{\lambda_1 e^{-\lambda_2 c} - \lambda_2 e^{-\lambda_1 c}}{\lambda_2 - \lambda_1}$$

$$F_{X+X}(c) = 1 - e^{-c\lambda}\,(1 + c\lambda)$$

Erlang-Verteilung: $f_{\sum_{i=1}^n X_i} = \frac{\lambda^n \cdot x^{n-1}}{(n-1)!} e^{-\lambda x}$

4.3.3. Normal- / Gaussverteilung $X \sim \mathcal{N}(\mu, \sigma^2)$

Modell für Verteilung von Messwerten, $W = \mathbb{R}$.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{(x-\mu)}{\sigma}\right)^2\right)$$

$E[X] = \mu \qquad \sigma_X = \sigma$
Verteilfunktion $F(x)$ nicht geschlossen darstellbar.

$X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), \quad X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2), \quad X_1 \pm X_2 \sim \mathcal{N}(\mu_1 \pm \mu_2, \sigma_1^2 + \sigma_2^2)$

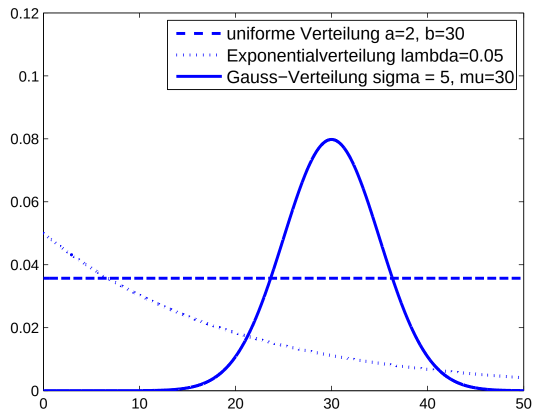
Standardnormalverteilung $\Phi = \mathcal{N}(0,1)$ ist tabelliert!
Transformation einer beliebigen Normalverteilung in die Standardnormalverteilung:

$$\boxed{F(x) = \mathbb{P}(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \Phi(u)}$$

$P(X \geq x) = 1 - \Phi(u), \qquad \Phi(u) = 1 - \Phi(-u)$
$P(a \leq X \leq b) = F(b) - F(a)$
$P(|X - \mu| \leq k \cdot \sigma) = 2\Phi(k) - 1$

4.3.4. Normalapproximation

$$P\left[\frac{s-\mu}{\sigma} \leq \frac{k-\mu}{\sigma}\right] = \Phi\left(\frac{k-\mu}{\sigma}\right) = P[s \leq K]$$



4.4. Log Normal $X \sim \mathcal{N}(\mu, \sigma^2)$

$Y = e^X$

$E[Y] = e\mu + \frac{\sigma^2}{2}$

$E[ln(Y)] = \mu$

4.5. Transformationen $Y = g(X)$

linear: $g(x) = a + b \cdot x, \quad \rightarrow \quad Y = a + b \cdot X$

$\mathbb{E}[Y] = a + b \cdot \mathbb{E}[X] \qquad Var(Y) = (|b|)^2 \cdot Var(X)$

$F_Y(y) = F_X(\frac{x-a}{b}), \quad$ falls $b > 0$

$F_Y(y) = 1 - F_X(\frac{x-a}{b}), \quad$ falls $b < 0$

$f_Y(x) = \frac{1}{|b|} f_x(\frac{x-a}{b})$

Standardisierung:

$g(x) = \frac{x-\mu}{\sigma}, \qquad \mathbb{E}[Z] = 0, \qquad Var(Z) = 1$

Allgemein *(monoton steigend):*

$$\mathbb{E}[Y] = \mathbb{E}[g(x)] = \int_{-\infty}^{\infty} g(x) \cdot f(x)dx \geq g(\mathbb{E}[X])$$

$$Var(Y) = E[Y^2] - E[Y]^2$$

$$f_Y(y) = F'_Y(y) = \left|\frac{1}{g'(g^{-1}(b))}\right| \cdot f_X(g^{-1}(y))$$

Quantile transformieren einfach mit: $q_{\alpha,X} \Rightarrow q_{\alpha,Y} = g(q_{\alpha,X})$

4.6. Simulation von Verteilungen

Sei $X \sim U(0,1)$ und $Y = F^{-1}(X)$ mit beliebigem F!
Y hat gerade die Verteilungsfunktion F.

1. $X \sim U(0,1) \Rightarrow A = \{x_1, x_2, \dots, x_n\}$
2. $F(Y) = F^{-1}(X) = F^{-1}(A)$

5. Mehrdimensionale Verteilungen

5.1. Diskret

Wahrscheinlichkeitsfunktion

$\mathbb{P}(X=x, Y=y) \to$ Tabelle

$$\mathbb{P}(X \leq x, Y \leq y) = \sum_{\substack{x_i \in W_x | x_i \leq x \\ y_i \in W_y | y_i \leq y}} \mathbb{P}(X = x_i, Y = y_i)$$

Randverteilungen: WK'funktionen

$$\mathbb{P}(X=x) = \sum_{y \in W_Y} \mathbb{P}(X=x, Y=y)$$

$$\mathbb{P}(Y=y) = \sum_{x \in W_X} \mathbb{P}(X=x, Y=y)$$

Zweidimensionale Verteilungstabelle

$$F(x,y) = \sum_{x_i \leq x} \sum_{y_k \leq y} p(x_i, y_k)$$

Bedingte Verteilungen: WK'funktionen

$$\mathbb{P}(X=x|Y=y) = \frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(Y=y)}$$

$$\mathbb{P}(Y=y|X=x) = \frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(X=x)}$$

Erwartungswert von $g(X,Y)$, $(g: \mathbb{R}^2 \to \mathbb{R})$

$$\mathbb{E}[g(X,Y)] = \sum_{\substack{x \in W_x \\ y \in W_y}} g(x,y) \mathbb{P}(X=x, Y=y)$$

Bedingter Erwartungswert

$$\mathbb{E}[X=x|Y=y] = \sum_{x \in W_x} x \cdot \mathbb{P}(X=x|Y=y)$$

$$\mathbb{E}[Y=y|X=x] = \sum_{y \in W_y} y \cdot \mathbb{P}(Y=y|X=x)$$

Unabhängigkeit zwischen X und Y $\Leftrightarrow$

$$\mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x)\mathbb{P}(Y=y), \forall x \in W_x, y \in W_y$$

Summe zweier unabhängiger ZV

$$P(X+Y \in [a,b]) = \sum_{n \in [a,b]} \sum_{k=0}^{n} P(X=k) \cdot P(Y=n-k)$$

5.2. Stetig

$$P[(X,Y) \in A] = \int \int_A f_{X,Y}(x,y) dx dy$$

$$F(x,y) = \mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v) du\, dv$$

Dichte:

$f(x,y), \qquad f: \mathbb{R}^2 \to \mathbb{R}$

Randverteilungen: Dichten

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx$$

Bedingte Verteilungen: Dichten

$$f_Y(y|X=x) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad f_X(x|Y=y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Unabhängigkeit von X und Y $\Leftrightarrow$**:**

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \qquad F(x_1, x_2) = F(x_1) \cdot F(x_2)$$

Erwartungswert von $g(X,Y)$, $(g: \mathbb{R}^2 \to \mathbb{R})$

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \cdot f_{X,Y}(x,y) dx dy$$

Bedingter Erwartungswert

$$E[X=x|Y=y] = \int_{-\infty}^{\infty} x f_X(x|Y=y) dx$$

$$E[Y=y|X=x] = \int_{-\infty}^{\infty} y f_Y(y|X=x) dy$$

5.3. Transformationen

Seien X und Y zwei **unabhängige** Zufallsvariablen mit Dichten f_X, f_Y.

Summe: $f_{X+Y}(t) = f_X(t) * f_Y(t) = \int_0^t f_X(u) \cdot f_Y(t-u)\, du$

$$P(X+Y \leq z) = \int_0^z \int_0^t f_X(u) \cdot f_Y(t-u) du dt$$

Produkt: $f_{X_1 \cdot X_2}(z) = \int_{X_1 \cdot X_2 = z} \frac{1}{|t|} f_{X_1}(t) \cdot f_{X_2}\left(\frac{z}{t}\right)\, dt$

Quotient: $f_{X_1/X_2}(z) = \int_{X_1/X_2 = z} |t| f_{X_1}(z \cdot t) f_{X_2}(t)\, dt$

Vorsicht: Oft abschnittsweise def. Funktionen!

5.4. Kovarianz und Korrelation

Kovarianz und Korrelation sind Kennzahlen, welche die Abhängigkeit von Zufallsvariabeln beschreiben.

Kovarianz:

$$\boxed{Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]}$$

Die Korrelation misst Stärke und Richtung der **linearen Abhängigkeit** zwischen X und Y.

$$\boxed{Corr(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}}$$

Rechenregeln:
$E[X+Y] = E[X] + E[Y]$ für beliegige, (abhängige) ZV
$Cov(X,Y) = E[XY] - E[Y] \cdot E[X]$
$Cov(X,Y) = 0$ falls X und Y unabhängig
$Cov(a+bX, c+dY) = b \cdot d \cdot Cov(X,Y)$
$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
$Cov(Z+X, Z+Y) = Cov(Z,Z) + Cov(Z,Y) + Cov(X,Z) + Cov(X,Y)$

$Corr(X,Y) = +1$ iff $Y = a+bX$ für $b > 0$
$Corr(X,Y) = -1$ iff $Y = a+bX$ für $b < 0$

Wenn X,Y **unabhängig**: $\Rightarrow Corr(X,Y) = 0$

5.5. Lineare Prognose

Lineare Prognose von Y gestützt mit Ansatz: $\widehat{Y} = a + bX$:

$$\widehat{Y} = \mu_Y + \frac{Cov(X,Y)}{Var(X)}(X - \mu_X)$$

$$E[(Y-\widehat{Y})^2] = (1-\rho_{XY}^2) \cdot Var(Y)$$

5.6. Zwei-dimensionale Normalverteilung

Kovarianz-Matrix:

$$\Sigma = \begin{pmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{pmatrix}$$

Normalverteilung:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \Sigma\right)$$

Dichte:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{det\Sigma}} \cdot \exp\left(-\frac{1}{2}(x-\mu_X, y-\mu_Y)\Sigma^{-1}\begin{pmatrix} x-\mu_X \\ y-\mu_Y \end{pmatrix}\right)$$

6. Grenzwertsätze

Die n-fache Wiederholung eines Zufallsexperimentes ist selber wieder ein Zufallexperiment.

6.1. independent and identically distributed

- Ergebnisse im ursprünglichen Experiment $A_1, ..., A_n$ sind unabhängig
- $P[A_1] = ... = P[A_n] = P[A]$ gleiche WK
- Zufallsvariabeln $X_1, ..., X_n$ sind **unabhängig**
- alle X_i haben **dieselbe Verteilung**

Mit der i.i.d. Annahme gilt:

$P[A \cap B] = P[A] \cdot P[B]$

$F_{x_1} = F_{x_n} \rightarrow P(x_1 \leqslant t) = P(x_n \leqslant t)$
$E[X_1 \cdot X_2] = E[X_1] \cdot E[X_2]$

6.2. Funktionen von Zufallsvariablen

Anstelle von $X_1, ...X_n$ werden neue Zufallsvariabeln als Funktion der alten gebildet.

- Summe: $S_n = X_1 + ... + X_n$
- arithmetisches Mittel: $\overline{X}_n = \frac{S_n}{n}$

Sonderflälle von S_n, bei denen die Bestimmung einfach ist:

1. Wenn $X_i \in \{0, 1\}$ (Bernoulliexperiment), dann ist $S_n \sim BI(n, p)$ mit $p = P[X_i = 1]$
2. Wenn $X_i \sim POI(\lambda)$, dann ist $S_n \sim POI(n \cdot \lambda)$
3. Wenn $X_i \sim N(\mu, \sigma^2)$, dann ist $S_n \sim \mathcal{N}(n \cdot \mu, n \cdot \sigma^2)$
4. Wenn $X_1, ..., X_n$ i.i.d und Normalverteilt mit $\mu_1, ..., \mu_n$ und $\sigma_1, ..., \sigma_n$, dann ist $S_n \sim \mathcal{N}(\sum \mu, \sum \sigma^2)$

Streuung der Summe wächst langsamer als n:

$E[S_n] = n \cdot E[X_i]$

$Var(S_n) = n \cdot Var(X_i)$

$\sigma_{S_n} = \sqrt{n} \cdot \sigma_{X_i}$
Streuung des arithm. Mittels nimmt ab:

$E[\overline{X}_n] = E[X_i]$

$Var(\overline{X}_n) = Var(X_i)/n$
Vergrösserung n führt zu „dünnerer“ Dichte.

$\lim_{n \to \infty} Var(\bar{X}_n) = 0$

6.3. Das Gesetz der Grossen Zahlen

(Moivre-Laplace)
$X_1, ..., X_n$ i.i.d mit μ

$\boxed{\overline{X}_n \rightarrow \mu \ (n \rightarrow \infty)}$

$\boxed{f_n[A] \rightarrow P[A] \ (n \rightarrow \infty)}$

6.4. Der Zentrale Grenzwertsatz

$X_1, ..., X_n$ i.i.d. mit μ und σ^2, dann gilt für grosse n:

$\boxed{S_n \approx \mathcal{N}(n \cdot \mu, n \cdot \sigma^2)}$

$\boxed{\overline{X}_n \approx \mathcal{N}(\mu, \sigma^2/n)}$

6.5. Chebychev Ungleichung

Die Konvergenz der Folge $(Y_1...Y_n)/n$ gegen $E(Y)$ erfolgt umsoschneller, je kleiner $Var(Y)$.

$$P[|\overline{X}_n - \mu| > c] \leqslant \frac{\sigma^2}{n \cdot c^2}$$

Mit dieser ist man stets auf der sicheren Seite, dafür aber meistens ziemlich grob.

7. Deskriptive Statistik

7.1. Kennzahlen

Arithmetisches Mittel, abhängig von Ausreissern, TR: mean(), $\bar{x}$:

$$\overline{x} = \frac{1}{n}(x_1 + ...x_n)$$

Empirische Varianz, TR: variance(), S_x^2:

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

Empirische Kovarianz:

$$s_{xy} = \frac{1}{n-1}\sum_{n-1}^{n}(x_i - \overline{X})(y_i - \overline{Y})$$

Empirischer Korrelationskoeffizient:

$$r = \frac{s_{xy}}{s_x \cdot s_y} \in [-1, 1], \qquad \begin{array}{l} r = +1 \Leftrightarrow y_i = a + bx_i, b > 0 \\ r = -1 \Leftrightarrow y_i = a + bx_i, b < 0 \end{array}$$

$sign(r)$ gibt Richtung,$|r|$ die Stärke der linearen Abhängigkeit.

Empirische kumulative Verteilungsfunktion:

$$F_n(x) = \frac{1}{n}\text{Anzahl}\{i | x_i \leq x\} \in [0, 1]$$

Empirisches α-Quantil q_α, $(0 < \alpha < 1)$:
Bei **geordneten Werten** $x_{(1)} \leqslant x_{(2)} \leqslant ... \leqslant x_{(n)}$ ist $q_\alpha = x_{(k)}$, wobei mit k die kleinste ganze Zahl, so dass $x_{(k)} > \alpha n$.
Die Werte werden also etwa im Verhältnis $\alpha : (1-\alpha)$ aufgeteilt.
Das Quantil ist *robust gegenüber Ausreissern*.

$$q_\alpha = \begin{cases} \frac{1}{2}\left(x_{(\alpha \cdot n)} + x_{(\alpha \cdot n + 1)}\right), & \text{if } n \text{ gerade} \\ x_{(\lceil \alpha \cdot n \rceil)}, & \text{else} \end{cases}$$

Median	$q_{0.5}$, TR: median()
unteres Quartil	$q_{0.25}$
oberes Quartil	$q_{0.75}$
Quartilsdifferenz	QD $= q_{0.75} - q_{0.25}$
	robuste Kennzahl für Streuung

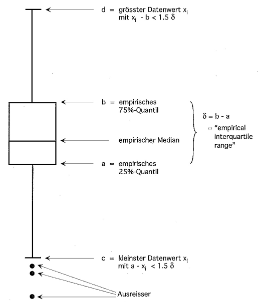
7.2. Histogramm, Boxplot und Q-Q Plot

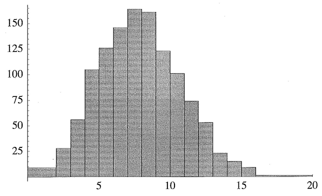
Histogramm: Einteilung der Werte in Klassen:

- Intervalle $(c_{k-1}, c_k]$
- h_k: Häufigkeit, Anzahl Werte im Intervall
- Fläche der Balken $\sim h_k$
- Höhe der Balken $\sim \frac{h_k}{c_k - c_{k-1}}$

Boxplot: Rechteck begrenzt durch das 25%- und 75%-Quantil und dickem Strich für den Median, Linien von grösstem bis kleinstem “normalen“ Wert (max. 1.5 mal die Quartilsdifferenz), Ausreisser: Sterne

Q-Q-Plot: “Quantil-Quantil-Plot“, es wird das theoretische Quantil gegenüber dem empirischen Quantil aufgetragen. Aussage möglich, wie gut die Daten dem Modell entsprechen. Bei Normalverteilung: Gerade mit Steigung σ und y-Achsenabschnitt μ.





8. Schliessende Statistik

Eine Beobachtung gegeben → plausibel?

8.1. Das Testproblem

Nullhypothese: $H_0 : p = p_0$

Alternativen:
$H_A : p \neq p_0$ (zweiseitig)
$H_A : p > p_0$ (einseitig nach oben, rechtsseitig)
$H_A : p < p_0$ (einseitig nach unten, linksseitig)

Signifikanzniveau: $P_{p_0}[X \geqslant c] \leqslant \alpha$

Fehler 1. Art: Verwerfen der Nullhypothese, obwohl sie richtig ist.

Fehler 2. Art: kein Verwerfen der Nullhypothese, obwohl sie falsch ist.

Beispiel: Nullhypothese: Mail ist kein Spam.
Fehler 1. Art: wichtiges E-Mail wird als Spam gekennzeichnet.
Fehler 2. Art: Spam wird als wichtiges Mail gekennzeichnet.

Ablauf eines Testes

1. Modell, e.g. $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2), \mu$ unbekannt
2. Nullhypothese H_0
3. Alternative H_A: zweiseitig / einseitig, oben / unten
4. Signifikanzniveau: $\alpha = 0.05$ oder 0.01
5. Verwerfungsbereich für H_0, so dass: $P[\text{Fehler 1.Art}] \leqslant \alpha$
6. Entscheidung

8.2. Binominalverteilt: $X \sim Bin(n, p)$

Fehler 1. Art einseitig, wenn H_0 für $x \geq c$ verworfen wird:

$$P_{p_0}[X \geqslant c] = \sum_{k=c}^{n} \binom{n}{k} p_0^k (1 - p_0)^{n-k}$$

Fehler 1. Art zweiseitig, wenn Nullhypothese für $c_2 \leq x \leq c_1$ verworfen wird:

$$\sum_{k=0}^{c_1} \binom{n}{k} p_0^k (1 - p_0)^{n-k} \leqslant \frac{\alpha}{2} \qquad \sum_{k=c_2}^{n} \binom{n}{k} p_0^k (1 - p_0)^{n-k} \leqslant \frac{\alpha}{2}$$

Wenn N gross ist, wird der Zentrale Grenzwertsatz benutzt: $X \sim N(n \cdot p, n \cdot p(1 - p))$
Fehler 2. Art

$$1 - \sum_{k=c}^{n} \binom{n}{k} p_0^k (1 - p_0)^{n-k}$$

8.3. P-Wert

P-Wert = **kleinstes Signifikanzniveau**, für welches H_0 verworfen wird (mit der vorliegenden Messung).

$p < \alpha \to H_0$ verwerfen

8.4. Vertrauens- /Konfidenzintervalle ($\alpha \downarrow, VI \uparrow$)

Im Vertrauensintervall zum Niveau $1 - \alpha$, wird eine Beobachtung als plausibel gewertet. Damit H_A bewiesen werden kann, muss $H_0 \notin$ VI sein.

Zweiseitiges $(1 - \alpha)$ Vertrauensintervall allgemein:

$$\hat{\theta} \pm \left(1 - \frac{\alpha}{2}\right)_{\text{Quantil}} \cdot \sqrt{Var(\hat{\theta})}$$

linksseitig ($\theta_A < \theta_0$): VI $\in \left(-\infty\,,\, \hat{\theta} + (1 - \alpha)_{\text{Quantil}} \cdot \sqrt{Var(\hat{\theta})}\right]$

rechtsseitig ($\theta_A > \theta_0$): VI $\in \left[\hat{\theta} - (1 - \alpha)_{\text{Quantil}} \cdot \sqrt{Var(\hat{\theta})}\,,\, \infty\right)$

8.4.1. Häufig gebrauchte Verteilungen, zweiseitig

Bem: einseitig analog, einfach $\frac{\alpha}{2} \to \alpha$ und Richtung anpassen
$X \sim Bin(n, p) \approx \mathcal{N}(np, np(1 - p))$, n gross, $\hat{\theta} = \hat{p}$:

$$\boxed{\hat{p} \pm \Phi^{-1}(1 - \frac{\alpha}{2}) \sqrt{\hat{p}(1 - \hat{p}) \frac{1}{n}}} \qquad \hat{p} = \frac{x}{n}, \quad \text{x: gemessene Erfolge}$$

Beachte: $Var(\hat{p}) = Var\left(\frac{x}{n}\right) = \frac{1}{n^2} Var(X) = \frac{\hat{p} \cdot (1 - \hat{p})}{n}$

$X \sim Poi(\lambda) \approx \mathcal{N}(\lambda, \lambda)$, $\hat{\theta} = \lambda$: $\boxed{\lambda \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\lambda}}$

$X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2), \sigma$ bekannt: $\boxed{\hat{x} \pm z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}}$

$X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2), \sigma$ unbekannt: $\boxed{\hat{x} \pm t_{(n-1, 1-\frac{\alpha}{2})} \cdot \frac{S_n}{\sqrt{n}}}$

Hypergeometrisch (Urnenmodell ohne Zurücklegen), n gross, N sehr gross: $\left[p \pm z\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{p(1-p)}{n}} \sqrt{\frac{N-n}{N-1}}\right]$

8.5. Macht eines Tests

Die Macht eines Testes ist die Wahrscheinlichkeit, dass dieser für ein $p \neq p_0$ richtig verwirft:

$p = 1 - \beta(\mu) = 1 - \mathbb{P}(\text{Fehler 2. Art})$

Je kleiner der Unterschied zwischen p_0 und p, desto kleiner die Macht und desto schwieriger ist es für den Test richtig zu entscheiden.

Wahrscheinlichkeit eines **Fehler 2. Art**:

$\beta(\mu) = P[\text{ Test akzeptiert } H_0 \text{ obschon ein } \mu \in H_A \text{ stimmt }]$

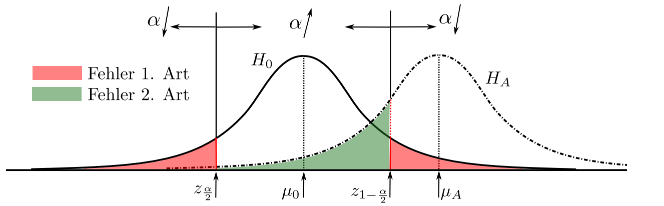
$X \sim Bin(n, p)$, $H_A \in [c, n]$ Dann ist die Macht: $1 - \beta(\mu) = P[\text{ Test verwirft } H_0 \text{ für} \mu \in H_A] = \sum_c^n \binom{n}{k} p^k \cdot (1 - p)^{n-k}$.

Allgemeines Vorgehen:

1. Verwerfungsbereich von H_0 bestimmen, z.B.

   $x \geq np_0 + \Phi^{-1}(1 - \alpha) \cdot \sqrt{np_0(1 - p_0)}$

2. $\mathbb{P}(\text{Fehler 2. Art}) = \beta = P_{H_A}(X \leq x) = \Phi\left(\frac{x - np_A}{\sqrt{np_A(1 - p_A)}}\right)$
3. Macht $p = 1 - \beta$



Bei zweiseitigen Tests müssen streng genommen beide Grenzen beachtet werden. Bsp. mit z-Test:

$P(\text{Fehler 2. Art}) = P_{HA}(X \leq c_o) - P_{HA}(X \leq c_u)$

$$= \Phi\left(\frac{\left[\mu_0 + \frac{\sigma_0}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}\right] - \mu_A}{\frac{\sigma_A}{\sqrt{n}}}\right) - \Phi\left(\frac{\left[\mu_0 - \frac{\sigma_0}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}\right] - \mu_A}{\frac{\sigma_A}{\sqrt{n}}}\right)$$

9. Statistik bei normalverteilten Daten

9.1. Punktschätzungen

$$\boxed{\widehat{\mu} = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i}$$

$$\boxed{\widehat{\sigma}^2 = S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \widehat{\mu})^2}$$

Erwartungswert der Schätzer:

$E(\widehat{\mu}) = \mu \quad E(\widehat{\sigma}^2) = \sigma^2$

9.2. z-Test

$X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ unbekannt, aber σ bekannt,

$$Z = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} \sim \mathcal{N}(0,1), \qquad z_{1-\alpha} = \Phi^{-1}(1-\alpha)$$

H_0 verwerfen, falls:
zweiseitig: $\boxed{|Z| > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}$
Zweiseitiger Verwerfungsbereich für $\bar{X}$:

$$\left(-\infty\,,\, \mu_0 - \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right] \cup \left[\mu_0 + \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\,,\, \infty\right)$$

einseitig:
$\mu_A > \mu_0$: $\boxed{Z > \Phi^{-1}(1-\alpha)}$ $\qquad \mu_A < \mu_0$: $\boxed{Z < \Phi^{-1}(\alpha)}$

9.3. t-Test

$X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ, σ unbekannt.

$$T = \frac{\bar{x} - \mu_0}{S_n}\sqrt{n} \sim t_{n-1}, \qquad \text{unter } H_0,\ n-1 \text{ FG}$$

H_0 verwerfen, falls:
zweiseitig: $\boxed{|T| > t_{n-1,1-\frac{\alpha}{2}}}$
Verwerfungsbereich für T: $\left(-\infty\,,\, -t_{n-1,1-\frac{\alpha}{2}}\right] \cup \left[t_{n-1,1-\frac{\alpha}{2}}\,,\, \infty\right)$
einseitig:
$\mu_A > \mu_0$: $\boxed{T > t_{n-1,1-\alpha}}$ $\qquad \mu_A < \mu_0$: $\boxed{T < -t_{n-1,1-\alpha}}$

Bem: Bei nicht normalverteilten Daten besteht die Gefahr auf einen sehr grossen Fehler 2. Art.

10. Punktschätzungen: allgemeine Methoden

Die Verteilung von X_i sei bekannt bis auf einen unbekannten Parameter θ, dabei kann θ auch mehrere Komponenten haben.

10.1. Momentenmethode

Unbekannte Parameter θ mit Hilfe der Momente $\boxed{\mu_k = E[X^k]}$ ausdrücken.

Binomial: $p = \theta$, Poisson: $\theta = \mu$, Standardnormalverteilung: $\theta = (\mu, \sigma^2)^T$

Vorgehen:

1. Unbekannte Parameter θ als Funktion der Momente $\mu_k = \mathbb{E}\left(X^k\right)$ schreiben
2. Ersetze wahre μ_k durch empirische Momente $\hat{\mu}_k$
$\hat{\theta}_j = g_j(\hat{\mu}_1, ..., \hat{\mu}_p)$ mit $\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$

10.2. Maximum-likelihood Schätzer

θ, so dass die log-Likelihood-Funktion maximiert wird:

- für diskrete X_i: $l(\theta) = \sum_{i=1}^{n} \ln(p_\theta(X_i))$
- für stetige X_i: $l(\theta) = \sum_{i=1}^{n} \ln(f_\theta(X_i))$

Vorgehen

- Likelihood-Funktion aufstellen: $L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta)...$
- Berechne $\hat{\theta}$, so $L(\theta)$ maximal wird, d.h. $\frac{dL}{d\theta} = 0$
- Hat θ mehrere Komponenten, so leite nach jedem einzeln ab.

Bernoulli: $\hat{p} = \frac{\sum x_i}{n}$, $\quad$ POI: $\lambda = \frac{\sum x_i}{n}$

10.3. Erwartungstreue

Ein Schätzer $\hat{\theta}$ ist auch eine Zufallsvariable und heisst **erwartungstreu**, wenn $\mathbb{E}\left(\hat{\theta}\right) =$ wahrer Parameterwert.
Bsp: Schätzer $\hat{\theta} = \bar{X}_n$ für μ bei $X \sim \mathcal{N}$

$$\mathbb{E}\left(\bar{X}_n\right) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i\right) = \frac{1}{n} \cdot n \cdot \mu = \mu$$

$\rightarrow$ erwartungstreu!
Bsp: $X \sim Bin(n,p)$, Schätzer $\hat{p} = \frac{x}{n}$ (x: gemessene Erfolge)

$$\mathbb{E}\left(\hat{p}\right) = \mathbb{E}\left(\frac{x}{n}\right) = \frac{1}{n}\mathbb{E}\left(x\right) = \frac{1}{n} \cdot np = p$$

$\rightarrow$ erwartungstreu!
Gegenbeispiel: $X \sim \mathcal{N}$:

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 = \ldots = \frac{n-1}{n}\sigma^2 \neq \sigma^2$$

ist nicht erwartungstreu!

11. Vergleich zweier Stichproben

11.1. Gepaarte und ungepaarte Stichproben

Randomisierung: Zufällig gewählte Reihenfolge der Versuche, verschiedene **Versuchseinheiten** unter zwei verschiedenen **Versuchbedingungen** ergeben eine **ungepaarte Stichprobe**. Einzelne Tests müssen nicht gleiche Stichprobengrösse haben.
Bsp: 2 Medikamente an verschiedenen Stichproben getestet.

Gepaarte Stichproben: beide Versuchsbedingungen an derselben Versuchseinheit getestet. Notwendigerweise müssen die beiden Stichprobengrössen gleich sein.
Bsp: 2 Reifentypen durch gleiche Fahrer getestet:

11.2. Gepaarte Vergleiche

Differenz innerhalb der Paare: $\boxed{u_i = x_i - y_i}$

Kein Unterschied zwischen Versuchsreihen: $E[U_i] = 0$
$H_0 : E(u_i) = 0 \quad H_A : E(u_i) =>< 0$

Verschiedene mögliche Tests:

- t-Test, wenn Normalverteilt, QQ-Plot
- Vorzeichen-Test, wenn beliebig verteilt
- Wilcoxon-Test, wenn symmetrisch verteilt, Histogramm

11.3. Vorzeichen-Test

Annahme: NUR i.i.d, keine normalverteilten Daten
Daten: $X_1, \ldots, X_n$ und $Z_i = X_i - \mu$
Alternativ: $X_1, \ldots, X_n$; $Y_1, \ldots, Y_n$ und $Z_i = X_i - Y_i$
Vorzeichen: $sign(Z_1), ...sign(Z_n)$

$Y \sim Bern : sign(Z_i) > 0 \rightarrow 1, \text{ else } 0$

Teststatistik: V=Anzahl pos. Beobachtungen, gemessen: Q

$$\sum Y = V \sim BIN(n,p) \quad \text{mit } p = \frac{Q}{n} \text{ und } H_0 : p_0 = 0.5$$

H_0 verwerfen, falls:

zweiseitig: $P(V \leq Q) \cup P(V \geq n-Q) \overset{p_0=0.5}{=} 2 \cdot P(V < Q) \leq \alpha$
einseitig:

mit $H_A : p_A < 0.5 : \quad P(V \leq Q) = \sum_{i=0}^{Q} Bin(n, 0.5, i) < \alpha$

mit $H_A : p_A > 0.5 : \quad P(V \geq n-Q) = \sum_{i=Q}^{n} Bin(n, 0.5, i) < \alpha$

Bem: Bin-Verteilung **nur** bei $p = 0.5$ symmetrisch!

11.4. Wilcoxon-Test

Kompromiss: setzt weniger voraus als t-Test, nützt Daten aber besser aus als Vorzeichen-Test.

1. Ränge bilden: $Rang(|U_i|) = k$
   $k = 1$ bedeutet, dass $|U_i|$ den kleinsten Wert unter $|U_1|...|U_n|$ hat.
   Wenn einzelne $|U_i|$ zusammenfallen, teilt man die Ränge auf.

2. V_i ist Indikator, ob U_i positiv ist: $V(U > 0) = 1$ und $V(U < 0) = 0$

3. Verwerfung der Nullhypothese falls W zu gross, zu klein oder beides ist (Tabellen).

$$\boxed{W = \sum_{i=1}^{n} Rang(|U_i|) \cdot V_i}$$

Eigenschaften :U hält das Niveau α exakt, falls F um 0 symmetrisch und x_i i.i.d ist.

Fehler 2.Art von t-Test ist oft viel grösser als Fehler 2.Art von U.

Wilcoxon-Test ist in der Praxis dem t- oder Vorzeichen-Test vorzuziehen, da Daten meist nicht Normalverteilt sind.

11.5. Ungepaart: Zwei-Stichproben Test

Ungepaarte Stichproben, unabhängige Zufallsvariablen
Annahmen: Normalverteilt mit gleicher Varianz, $\sigma_X = \sigma_Y$

$$\left.\begin{array}{ll} X_i \overset{iid}{\sim} \mathcal{N}(\mu_X, \sigma^2), & i = 1, \ldots, n \\ Y_i \overset{iid}{\sim} \mathcal{N}(\mu_Y, \sigma^2), & j = 1, \ldots, m \end{array}\right\} \mu_X, \mu_Y, \sigma unbekannt$$

Nullhypothese $H_0 : \mu_X = \mu_Y$ verwerfen, falls:

Für $H_A : \mu_X \neq \mu_Y$: $|T| = \dfrac{|\bar{X}_n - \bar{Y}_m|}{S_{pool}\sqrt{1/n + 1/m}} > t_{n+m-2, 1-\frac{\alpha}{2}}$

Für $H_A : \mu_X > \mu_Y$: $T = \dfrac{\bar{X}_n - \bar{Y}_m}{S_{pool}\sqrt{1/n + 1/m}} > t_{n+m-2, 1-\alpha}$

Für $H_A : \mu_X < \mu_Y$: $T = \dfrac{\bar{X}_n - \bar{Y}_m}{S_{pool}\sqrt{1/n + 1/m}} < -t_{n+m-2, 1-\alpha}$

Bem: $\mu_{X0} - \mu_{Y0}$ jeweils weggelassen, da immer = 0.

$$S_{pool}^2 = \frac{1}{n+m-2}\left(\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 + \sum_{i=1}^{n}(Y_i - \bar{Y}_m)^2\right)$$
$$= \frac{1}{n+m-2}\left((n_X - 1) \cdot S_X^2 + (n_Y - 1) \cdot S_Y^2\right)$$

$$Var(\bar{X}_n - \bar{Y}_m) = \sigma^2(\frac{1}{n} + \frac{1}{m})$$

11.5.1. Vertrauensintervalle für $d = \mu_X - \mu_Y$

H_A wird nicht verworfen, falls $d \in VI$:
$H_A : \mu_X \neq \mu_Y$:

$$VI = \bar{X}_n - \bar{Y}_m \pm S_{pool}\sqrt{\frac{1}{n} + \frac{1}{m}} \cdot t_{n+m-2, 1-\frac{\alpha}{2}}$$

$H_A : \mu_X > \mu_Y$:

$$VI = \left(\bar{X}_n - \bar{Y}_m - S_{pool}\sqrt{\frac{1}{n} + \frac{1}{m}} \cdot t_{n+m-2, 1-\alpha}\, , \, \infty\right)$$

$H_A : \mu_X < \mu_Y$:

$$VI = \left(-\infty\, , \, \bar{X}_n - \bar{Y}_m + S_{pool}\sqrt{\frac{1}{n} + \frac{1}{m}} \cdot t_{n+m-2, 1-\alpha}\right)$$

A. Diverse Zusammenhänge

A.1. Korrelation - Unabhängigkeit

korreliert	$\Rightarrow$	abhängig
abhängig	$\not\Rightarrow$	korreliert
unabhängig	$\Rightarrow$	unkorreliert
unkorreliert	$\not\Rightarrow$	unabhängig
korreliert	$\Leftrightarrow$	Kovarianz $\neq 0$
unkorreliert	$\Leftrightarrow$	Kovarianz $= 0$
Korrelation $= 0$	$\Leftrightarrow$	$\text{Cov}(X, Y) = 0$

Ausnahme: Für 2D-Normalverteilung:

abhängig $\Leftrightarrow$ korreliert

Disclaimer

Diese Formelsammlung basiert auf der Formelsammlung von Andrea Fuchs.
Sie kann gerne verändert werden und darf unter Angabe aller bisherigen Autoren auch wieder veröffentlicht werden.

Es wird keine Garantie für Richtigkeit der angegeben Daten erteilt.
Verändert durch:
HS 09: Till Richter
HS 09: Bastian Wohlfender
HS 12: Cédric de Crousaz