Statistik

UE 1

Zufallsvorgang – mind. 2 mögl. Ergebnisse, die sich ausschließen, im voraus nicht bestimmbar, welches eintritt – Würfel & Wetter

Zufallsexperiment – Zufallsvorgang unter kontrollierten Bed., beliebig oft wiederholbar – Würfel ja, Wetter nein

Ereignis – Zusammenfassung mehrerer Ergebnisse – wenn Ereignis A eintritt war Exp. „erfolgreich"

Statistische Wahrscheinlichkeit

= „im Mittel erwartete Anzahl von A bei n Versuchen"
→ Schätzung, Empirisch
→ kann nur durch Beobachtung approximiert werden, nie berechnet!

Ereignis A, n Versuche

Erfolge $h_n(A)$ absolute Häufigkeit

relative Häufigkeit $f_n(A) = \frac{h_n(A)}{n} = \frac{\text{absolute H.}}{\text{Anzahl Versuche}}$

bei ∞ n

$$\boxed{\lim_{n \to \infty} f_n(A) = p}$$ statistische Wahrscheinlichkeit

Klassische Wahrscheinlichkeit

= La Place - Wahrscheinlichkeit

= Gedankenexperiment!
→ Voraussetz.:
- endlich viele Versuchsausgänge
- alle haben gleiche „Chance"
  ↳ „gleich wahrscheinlich"

$$\boxed{P(A) = \frac{\text{Anzahl günstige Ergebnisse}}{\text{Anzahl mögliche Ergebnisse}}}$$ = LaPlace-Experiment

Wahrscheinlichkeit eines Ergebnisses = $\frac{1}{N}$ (= p „probability")

Multiplikationssatz d. Kombinatorik: mehrstufiges Exp. Anzahl mögl. Ergebnisse = mögliche 1. Stufe × (dann noch) mögl. 2. Stufe
(Würfel 3x: 6 × 6 × 6)
(Urne 5 Kugeln, 3x ziehen: 5 × 4 × 3)

Mengen

disjunkt – kein gemeinsames Element ($A \cap B = \emptyset$) nur innere Punkte! Grenze kann gleich sein!

Intervalle: (a ; b]

offen – ausschließlich, ohne a

geschlossen, einschließl., incl. b

$x \in A$ – x ist Element A

$B \subset A$ – B ist Teilmenge von A

$B \subsetneq A$ – ", aber nicht gleich A „echte Teilmenge"

Rechenregeln für Mengen

- Kommutativ- / Assoziativgesetz für nur „∩" oder „∪" ganz „normal"
- Distributivgesetz $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
  $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
- De Morgan $\overline{A \cup B} = \overline{A} \cap \overline{B}$ „Komplemente aufteilen indem Zeichen rumdrehen"
  $\overline{A \cap B} = \overline{A} \cup \overline{B}$
- Mengendifferenz $A \setminus B = A \cap \overline{B} = (A \cup B) \setminus B = A \setminus (A \cap B)$
  gilt auch für Wahrscheinlichkeiten

Rechenregeln Wahrscheinlichkeit

Ergebnisraum (= Ereignisraum, Grundraum, Stichprobenraum)
= Menge aller möglichen Ergebnisse eines Zufallsvorganges Ω $P(\Omega) = 1$ → Ereignis = Teilmenge
↳ sicheres Ereignis, $\emptyset$ unmögliches Ereignis

wenn A und B disjunkt, dann $P(A \cup B) = P(A) + P(B)$
wenn nicht: $= P(A) + P(B) - P(A \cap B)$
„Additionsaxiom"

$P(\Omega) = 1$ „Normierungsaxiom"

$P(A) \geq 0$

Axiome von Kolmogoroff

Wahrscheinlichkeitsmaß = eine Funktion, die Teilmengen einer Menge (unter Einhaltung der Axiome) Zahlen (Wahrscheinlichkeiten bzw. 0 und 1) zuordnet.

Bedingte Wahrscheinlichkeit

= klassische Wahrscheinlichkeit mit eingeschränktem Ergebnisraum

P(A) und P(B) größer 0!

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Wahrscheinlichkeit, dass beide eintreten

Wahrscheinlichkeit schon eingetretenes

„Wahrscheinlichkeit von A bedingt auf B" (Wahrsch. von A wenn B schon eingetreten ist)

bzw. Produktsatz

$P(B) > 0$ $P(A \cap B) = P(A \mid B) \cdot P(B)$

$P(A) > 0$ $P(A \cap B) = P(B \mid A) \cdot P(A)$

mehr Ergebnisse bzw. Stufen $P(A \cap B \cap C) = P(A \mid B \cap C) \cdot P(B \mid C) \cdot P(C)$

Unabhängigkeit

= Wahrscheinlichkeit von A hängt nicht vom Eintreten von B ab.

zB. Würfeln (Würfel hat kein Gedächtnis)
+ Urne mit zurücklegen
nicht: ohne "

$P(A|B) = P(A)$ $P(B) \neq 0$

oder

$P(A \cap B) = P(A) \cdot P(B)$ weil dann disjunkt

→ wenn A & B unabhängig sind sämtliche daraus durch beliebige Operationen konstruierte Ereignisse auch unabhängig

→ Unabhängigkeit mehr als 2 Ereignisse:

$$P(A_1 \cap A_2 \cap \ldots \cap A_i) = P(A_1) \cdot P(A_2) \cdot \ldots \cdot P(A_n)$$

Totale Wahrscheinlichkeit

A_{1-n} disjunkte Zerlegung von Ω

Wahrscheinlichkeit Alternative 1 fehlerhaft x Anteil am Gesamt
\+ " A 2 " x "
\+ A 3 x "
... = totale Wahrsch.

auch: $P(B) = \sum_{i=1}^{k} P(B|A_i) \times P(A_i)$

Formel von Bayes

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{i=1}^{k} P(B|A_i) P(A_i)}$$

$\frac{\text{totale Wahrsch. eines Alternative}}{\text{" " aller A}}$

Wahrscheinlichkeit einer Alternative wenn sicher, dass fehlerhaft

---

Notation: $P(\{1\})$ weil Teilmenge
$P(\{(2,3)\}) = 1/36$ $P(\{2,3\}) = 1/3$

VL 2

Diskrete ZV

$P(X=x)$

Zufallsvariable (ZV): Variable, deren Ausprägungen Ergebnisse eines Zufallsvorganges sind (reelle Zahlen)

Realisierung/Wert (von X): konkreter Zahlenwert, den ZV bei Durchführung annimmt $P(X=x)$

diskret: wenn ZV nur endlich / abzählbar ∞ Werte annehmen kann zB $\mathbb{N}$

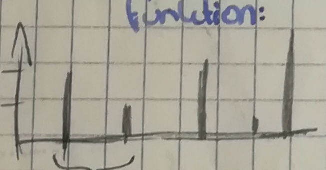
Einzelwahrscheinlichkeit: Wahrscheinlichkeit, dass ZV speziellen Wert annimmt $P(X=x)$

Träger: Menge aller Werte, die positive Wahrsch. haben (möglicherweise eintreffen)
$T_x = \{x : P(X=x) > 0\}$ $|T_x| = k \rightarrow$ k-punkt verteilt
1-punkt verteilt: uneigentl. ZV

Verteilungstabelle:

x (Wert)	.... alle Werte aus Träger ...
$P(X=x)$	Wahrsch, dass sie eintreffen

Wahrscheinlichkeitsfunktion: $\boxed{f(x)=P(X=x)}$ ordnet jedem Element im Träger die Wahrsch. zu mit der es eintrifft, allen anderen die null



oft mit Stabdiagrammen veranschaulicht
Summe aller Stäbe muss 1 sein

Wahrscheinlichkeit eines Intervalls: bei diskreten ZV: aufsummieren Einzelwahrsch.

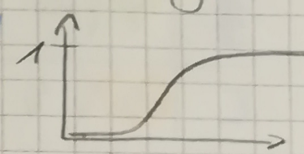
Stetige ZV

Intervallwahrscheinlichkeit = Wahrscheinlichkeitsverteilung = Verteilung der ZV
↳ Wahrscheinlichkeitsmaß P_x, das jedem Intervall $I \subset \mathbb{R}$ den Wert $P_x(I) = P(X \in I)$ zuordnet.
Wahrsch, dass X in I liegen wird

Verteilungsfunktion VF $\boxed{F(x) = P(X \leq x)}$ auch F_x
↳ ordnet jeder reellen Zahl die Wahrsch. zu, dass X kleiner oder gleich x sein wird
Werte immer zw. 0 & 1!

Eigenschaften Verteilungsfunktion:

- $P(a < X \leq b) = P(\{X \leq b\} \setminus \{X \leq a\}) = F(b) - F(a)$
  $a < b$! Wahrsch., dass X in $(a, b]$ fällt, x größer a bis b
- monoton wachsend
- rechtsstetig (wenn Sprünge)
  (der Sprungstelle wird der rechte (höhere) Wert zugeordnet
- $\lim_{x \to -\infty} F(x) = 0$ / $\lim_{x \to \infty} F(x) = 1$



- bei diskreten ZV: Treppenform

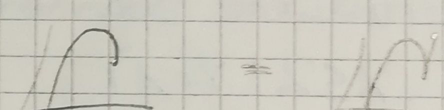
(Wahrscheinlichkeits-)Dichte:

Funktion $f(x)$, sodass für alle reellen Zahlen $a \leq b$

$$\boxed{P(a \leq X \leq b) = \int_a^b f(x)\,dx}$$

→ Intervallwahrsch. = Flächeninhalt unter Dichte von a bis b!

- $f(x) \geq 0$ → muss aber nicht zw. 0 und 1 liegen!
- $\int_{-\infty}^{\infty} f(x)\,dx = 1$



- „fast überall" eindeutig bestimmt wenn sie sich an einzelnen Stellen unterscheiden (endlich viele) sind alle Integrale gleich, da Linien keine Fläche haben → beschreiben gleiche Verteilung
- $P(X = x) = 0$ alle Einzelwahrsch. sind gleich null!

Verteilungsfunktion eines stetigen ZV ist stetig! (keine Sprünge!)

L_D = Dichtefunktion !?! ableiten

Zufallsvektoren

geordnet (X, Y); $P(X = x, Y = y)$

- (gemeinsame) Wahrscheinlichkeitsfunktion $f(x,y) = f_{xy}(x,y)$
- Kontingenztabelle

Verteilung von X

X \ Y	0	1	Σ
1			
2			
3			
Σ			1

Verteilung von Y

→ $P(X = 1, Y = 0)$
→ $P(X = 2, Y \text{ egal})$
→ $P(X \text{ egal}, Y = 1)$

Randverteilungen zu X & Y

Bedingte Wahrscheinlichkeit (für zweidimensionale diskrete Zufallsvektoren)

$$P(X=x \mid Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{f_{xy}(x,y)}{f_y(y)}$$

X\Y	1	2	3	Σ
1	a	d	g	1/4
2	b	e	h	1/2
3	c	f	i	1/4
Σ	1/2	1/4	1/4	1

- $P(X=1 \mid Y=1) = \frac{a}{1/2}$
- $P(X=2 \mid Y=1) = \frac{b}{1/2}$
- $P(Y=3 \mid X=2) = \frac{h}{1/2}$
- $P(Y=1 \mid X=2) = b/{1/2}$

„Spaltenwerte durch Spaltensumme"

„Zeilenwerte durch Zeilensumme"

Bedingte Wahrscheinlichkeitsfunktionen von x gegeben y

$$f_{X|Y}(x|y) = P(X=x \mid Y=y) = \frac{f_{x,y}(x,y)}{f_y(y)}$$

Bedingte Intervallwahrscheinlichkeit diskret

↳ Einzelwerte zusammenaddieren ! alle aufs gleiche Beding.!

zB. $P(X \leq 2 \mid Y=1)$

Erwartungswert

$E(X) = EX = \mu$ ('mü')

= Zahl zur Beschreibung der mittleren Lage einer ZV „Mittelwert ZV" → was ist „zu erwarten?"

- bei diskreter ZV: $E(X) = \sum a \cdot f(a) = \sum a \cdot p$

  a: alle Realisierungen, p: Einzelwahrsch.

bzw. $EX = \mu$ = Realisierung x deren Wahrsch. + nächste R. x W. + ..

- Funktion heißt symmetrisch bzgl. c wenn an c (EX) gespiegelt werden kann (f(c+x) = f(c−x)

  zwingend $c = EX$!

- bei stetiger ZV $EX = \int_{-\infty}^{\infty} x \cdot f_x(x) \cdot dx$

  - $f_x(X)$ = Dichte von X
  - wenn Integral endlich
  - → Integral unendlich, EX existiert nicht!

Indikatorfunktionen

verkürzte Bez. für Funktionen, die auf versch. Intervallen unterschiedl. definiert sind

zB. $f(x) = \begin{cases} 1 & \text{wenn } 0 \leq x < 0{,}5 \\ 2 & \text{wenn } 0{,}5 \leq x < 1 \\ x-1 & \text{wenn } 1 \leq x < \infty \\ 0 & \text{sonst} \end{cases}$

Fallunterscheidung

$f(x) = 1\, I_{[0,0{,}5)} + 2\, I_{[0{,}5,1)} + (x-1)\, I_{[1,\infty)}$

Indikatorfunktion

Integrale berechnen

$$\int_a^b f(x)\,dx = \Big[\underbrace{F(x)}_{\text{Stammfunktion}}\Big]_a^b = F(b) - F(a)$$

b: 2. Grenze
a: 1. Grenze

Stammfunktion bilden

- Hochzahl um eins erhöhen
- Faktor durch neue Hochzahl teilen
- „+ c"

→ wenn Funktion stückweise definiert, für jeden Abschnitt eigenes Integral!

Rechenregeln Erwartungswert

- **EX einer Funktion einer ZV**

1 diskrete ZV: alle möglichen Ergebnisse x in Funktion einsetzen & mit Einzelwahrsch. des x multiplizieren, dann alle addieren

zwingend auch diskret → zB. $g(x) = x^2 + 1$

x	1	2	3
P	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

$$E(g(x)) = (1^2+1)\cdot\tfrac{1}{3} + (2^2+1)\cdot\tfrac{1}{6} + (3^2+1)\cdot\tfrac{1}{2} = \tfrac{2}{3} + \tfrac{5}{6} + 5 = 6{,}5$$

2 stetige ZV: $E(g(x)) = \int_{-\infty}^{\infty} g(x) \cdot \underbrace{f_x(x)}_{\text{Dichtefunktion}}\,dx$

→ funktioniert analog auch mit Zufallsvektoren!

- **lineare Transformation**

wenn $Y = b + c \cdot X$ und ich kenne $E(X)$, ist $E(Y) = b + c \cdot E(X)$

- **Additivität**

1 $E(X+Y) = E(X) + E(Y)$ 2 $E(c \cdot X) = c \cdot E(X)$

- **Monotonie**

1 $X \leq Y \rightarrow E(X) \leq E(Y)$ 2 $E(X) = 0$ wenn $X \geq 0$ nur dann wenn X nur null werden kann ($P(X=0) = 1$)

- **Produktregel**

$E(X \cdot Y) = E(X) \cdot E(Y)$! nur für unabhängige ZV !

Varianz

$Var(X) \geq 0$ (0 nur bei Ein-Punkt-Verteilung) $= \sigma^2$ ("Sigma²")
= Streuung (Kenngröße, wie weit Verteilung ZV gestreut ist)

$Var(X) = E(X^2) - (E(X))^2$ Verschiebungsregel

→ diskrete ZV: $\sum_j (a_j - EX)^2 \cdot f(a_j)$ f = Wahrscheinlichkeitsfunktion

→ stetige ZV: $\int_{-\infty}^{\infty} (x - EX)^2 \cdot f(x)\, dx$ f = Dichtefunktion

- lineare Transformation: $Var(Y) = c^2 \, Var(X)$ $Y = b + c \cdot X$
- für ! unabhängige ! ZV: $Var(X+Y) = Var(X) + Var(Y)$

Standardabweichung $\sigma = \sqrt{Var(X)}$

Quantile

q_α (α zw. 0 & 1) → "Quantilsniveau" = α-Quantil zu X wenn $P(X \leq q_\alpha) = \alpha$

z.B. $q_{0,5}$ → Median
wenn genau die Hälfte der Dichte-fläche vor bzw nach q kommt d.h. wenn ein Ergebnis kleiner gleich q genauso wahrscheinlich ist wie ein Ergebnis größer q

Berechnung durch $F(x) = \alpha$ F(x) = Verteilungsfunktion ! bei stetig !
→ kann mehrere Lsg haben!

Lsg = inverse Funktion = $F^{-1}(\alpha)$ nicht: $1/F(\alpha)$!
⟺ Funktion aller Lsg., Funktion, die Verteilungsfunktion „aufhebt"

typisches Wertbereich bei unendlichen Trägern

$P(q_{\alpha/2} \leq X \leq q_{1-\alpha/2}) = 1 - \alpha$ (bei kleinem α)

finanzmathemat.: Quantile = „Value at risk"

Bernoulli-Verteilung

- $X = \begin{cases} 1 & A \text{ tritt ein} \\ 0 & A \text{ tritt nicht ein} \end{cases}$ $P(A) = \pi$ (Parameter) $T_X = \{0,1\}$

↳ kann nur diese beiden Ergebnisse haben!

- Wahrscheinlichkeitsfunktion:

$$f(x) = \pi \ \{\text{wenn } x = 1$$
$$f(x) = 1-\pi \ \{\text{wenn } x = 0, \text{ sonst } 0$$

→ Spezialfall Binomialverteilung ($n=1$)

- $EX = \pi$ $Var X = \pi \cdot (1-\pi)$

Binomialkoeffizient

„n über k"

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

→ wie viele Mengen mit k Elementen aus n möglich?

⚠ $0! = 1! = 1$

$$\binom{n}{n} = \binom{n}{0} = 1$$
$$\binom{n}{1} = n$$

Binomialverteilung

- n Versuche, Ereignis A mit $P(A) = \pi$ betrachtet
- X = erfolgreiche Versuche, $T_X = \{0, 1, \dots, n\}$
- $\boxed{X \sim B(n, \pi)}$ (X ist binomialverteilt mit n Versuchen und $P(A) = \pi$)
- Wahrscheinlichkeitsfunktion:

$$f(x) = \binom{n}{x} \cdot \pi^x (1-\pi)^{n-x} \quad x = 0, 1, \dots, n, \text{ sonst } 0$$

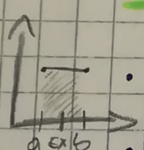
→ $\pi = 0{,}5$: symmetrisch, π größer linksschief, π kleiner rechtsschief

Poisson-Verteilung
- bei Daten zählen, unbekanntes n
- zählen wie oft X in festem Zeitraum auftritt

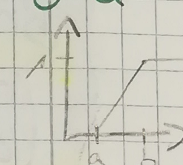
→ Grenzfall Binomialverteilung für $B(n = \text{sehr groß}, \pi = \text{sehr klein})$
langer Zeitraum, unwahrsch.
z.B. $n = 30$ $\lambda = 0{,}05$

- $\boxed{X \sim Po(\lambda)}$ $\lambda > 0$
- Wahrscheinlichkeitsfunktion:
$$f(x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda}$$
$x = \{0, 1, 2, \dots\}$ $T_x = \{0, 1, 2, \dots\}$! immer rechtsschief!
→ Träger nach oben unbegrenzt
- $E(X) = \underbrace{\lambda}_{\text{Lambda}}$, $= Var(X) = \lambda$
- für unabhängige X & Y: $X \sim Po(\lambda_1)$, $Y \sim Po(\lambda_2)$ → $X + Y \sim Po(\lambda_1 + \lambda_2)$

Stetige Gleichverteilung
Werte auf Intervall $[a, b]$ gleichmäßig verteilt

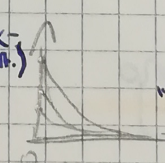


- Dichte $f(x) = \frac{1}{b-a}$ $a \leq x \leq b$ Verteilungsf. $F(x) = \frac{x-a}{b-a}$
- $E(X) = \frac{a+b}{2}$ $Var(X) = \frac{(b-a)^2}{12}$
- $X \sim G(a, b)$



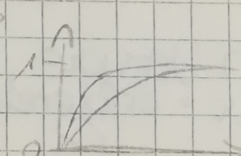
Exponentialverteilung
Lebensdauern, Wartezeiten: beliebige positive Werte

- $X \sim Exp(\lambda)$ $X \sim Exp(1)$ (Standardexponentialvert.)
- Dichte $f(x) = \lambda e^{-\lambda x}$ $x \geq 0$



„rechtsschief"

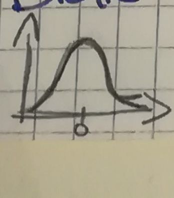
- Verteilungsf. $F(x) = (1 - e^{-\lambda x})$ $x \geq 0$

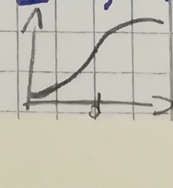


- $E(X) = \frac{1}{\lambda}$ $Var(X) = \frac{1}{\lambda^2}$

Normalverteilung

- $X \sim N(\mu, \sigma^2)$
- Dichte $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ Verteilungsfunktion ?
- $X \sim N(0, 1)$ Standardnormalverteilung
- Dichte $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ Verteilungsfunktion $\Phi(x)$ Tabelle!





Dichtetransformationsformel (linear)

wenn $Y = g(X) = b + c \cdot X$ $(c \neq 0)$ X ist stetige ZV mit Dichte f_X,
dann ist Y auch stetig mit Dichte $f_Y(y) = \frac{1}{|c|} \cdot f_X\left(\frac{y-b}{c}\right)$

→ wenn $X \sim N(\mu, \sigma^2)$ ist Y auch normalverteilt:
$$Y \sim N(b + c \cdot \mu, c^2 \sigma^2)$$

Standardisierung von ZV:

standardisiert wenn $\underline{E(X) = 0,\ Var(X) = 1}$

beliebige ZV $(EX = \mu, \sigma > 0)$ standardisiert durch $\underline{Z = \frac{X-\mu}{\sigma}}$

→ für $X \sim N(\mu, \sigma^2)$:
$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \Phi(z)$$

Quantile Standardnormalverteilung z_α

$\Phi(z_\alpha) = \alpha$ $\qquad$ $z_\alpha = -z_{1-\alpha}$ weil Dichte symmetrisch
$\qquad$ $z_\alpha = \Phi^{-1}(\alpha)$

→ für alle $X \sim N(\mu, \sigma^2)$: $\underline{q_\alpha = \mu + \sigma \cdot z_\alpha}$

- Summen unabhängiger $X \sim N(\mu, \sigma^2)$ oder $X \sim$ " $\cdot\, c = Y$
sind wieder normal verteilt.

Kovarianz

$Cov(X,Y) = \sigma_{XY}$ mit μ_X & μ_Y

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

Spezialfälle:
- $X = Y \to Cov = Var(X)$
- X & Y unabhängig $\to Cov = 0$

↳ positiv wenn gleichläufig, negativ wenn gegenläufig

Rechenregeln:
- $Cov(b + c \cdot X, a + d \cdot Y) = c \cdot d \cdot Cov(X,Y)$
- $Cov(X + Y, U) = Cov(X, U) + Cov(Y, U)$

Varianz: $Var(a \cdot X + b \cdot Y) = a^2\, Var(X) + b^2\, Var(Y) + 2 \cdot a \cdot b \cdot Cov(X,Y)$
auch wenn X & Y nicht unabhängig

Korrelation

Normierung der Kovarianz zur besseren Interpretation / Vergleich.

Korrelationskoeffizient $\rho_{XY} = \rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}} = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y}$

- liegt zw. -1 und 1
- X & Y unabhängig → wieder 0
- wenn $Y = b + c \cdot X$ $\quad c \neq 0$ dann $\rho(X,Y) = \begin{cases} 1 & \text{wenn } c > 0 \\ -1 & \text{wenn } c < 0 \end{cases}$
  → perfekter Zusammenhang
- unabhängig → unkorreliert (nicht umgekehrt da Korrelation „nur" lineares Zusammenhangsmaß, zeigt bei zB. quadratischen auch 0)

Grenzwertsätze

wenn $n \to \infty$, dann Approx.

Folgen von ZV: wenn jeder natürlichen Zahl eine ZV zugeordnet werden kann (zB. $\bar{X}_n$ oder $Y_n \in X_1 + X_n$)

unabhängig und identisch verteilt (u.i.v.) Zufallsvorgang, der unabhängig voneinander mehrmals durchgeführt wird
↳ alle Kennwerte gleich!

Gesetz d. großen Zahlen

$Z_n \in (c-\varepsilon, c+\varepsilon)$ „Epsilon-Umgebung" (während dann liegt Z_n gegen c)

→ für große n konvergieren Z_n gegen ein reelles c
→ arithmetisches Mittel konvergiert gegen EX

Ungleichung von Tschebyscheff

→ relative Häufigkeit gegen klassische Wahrsch.

$$P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}$$

Zentraler Grenzwertsatz

„Summen / arithmet. Mittel nicht normalverteilter ZV sind approx. normalverteilt"

approximativ: $E(X_1 + \ldots + X_n) = n \cdot \mu \quad | \quad Var(X_1 + \ldots + X_n) = n \cdot \sigma^2$
egal wie X_n verteilt

„jede Verteilungsfunktion konvergiert gegen die der Standardnormalverteilung $\Phi(z)$"

" nach Lindeberg-Levy

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n \cdot \sigma^2}} = \frac{\frac{1}{n}\sum_{i=1}^{n} X - \mu}{\sqrt{\sigma^2/n}} = \sqrt{n}\,\frac{\bar{X}_n - \mu}{\sigma}$$

$\bar{X}_n$ = arithmetisches Mittel
$E(\bar{X}_n) = \mu$
$Var(\bar{X}_n) = \frac{\sigma^2}{n}$ wenn X_i u.i.v mit $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$
↳ strebt gegen 0

nach Moivre

$$Z_n = \frac{H_n - n\cdot\pi}{\sqrt{n\cdot\pi\cdot(1-\pi)}} \overset{\text{approx.}}{\sim} N(0,1) \qquad EX = \pi \qquad Var X = n(1-\pi)$$

wenn X_i Bernoulliverteilt! $X \sim B(1,\pi)$

nicht für Einzelwahrsch.

Approx. Binomialverteilung mit Stetigkeitskorrektur

- wenn $H_n \sim B(n,\pi)$
- wenn $n\cdot\pi \geq 5$ und $n(1-\pi) \geq 5$
- a und b ganzzahlig

$$\underset{P(X=c)}{P(a \leq H_n \leq b)} \approx \Phi\left(\frac{\overset{c}{b} + 0{,}5 - \overbrace{n\cdot\pi}^{\mu}}{\sqrt{n\cdot\pi(1-\pi)}}\right) - \Phi\left(\frac{\overset{c}{a} - 0{,}5 - \overbrace{n\cdot\pi}^{\mu}}{\sqrt{n\cdot\pi(1-\pi)}}\right)$$

wenn $a \to \infty$ geht das $\to 0$

Punktschätzung

→ Näherungswert für unbekannte Parameter aus Daten bestimmen

- Daten → Statistisches Modell annehmen, zB. $X \sim N(\mu, \sigma^2)$ → unbekannten Parameter als θ („Theta") bezeichnen → Schätzer für θ finden → Werte ermitteln
- Grundgesamtheitsparameter: ein uns interessierender Parameter/Kennwert θ
- Schätzer/Schätzstatistik/Punktschätzer: ZV, Formel/Funktion der Stichprobenvar, approximativ
  ↳ darf Unbekannte nicht enthalten, nur Z_n
- Schätzwerte: Zahlen als Realisierungen des Schätzers
- Schätzer sollte: im Mittel richtigen Wert liefern, weniger um wahren Wert streuen wenn mehr Daten

Erwartungstreue

„macht keine Fehler"

- $T = g(X_1, \dots, X_n)$ erwartungstreu wenn $E_\theta(T) = \theta$
  bzw. $\underbrace{Bias_\theta(T)}_{\text{Verzerrung des Teststatistik}} = E_\theta(T) - \theta = 0$
- asymptotisch erwartungstreu wenn $E_\theta(T_n) \to \theta$ wenn $n \to \infty$

Beispiele
- arithmetisches Mittel $\frac{1}{n}\sum_{i=1}^{n} E(X_i)$ für $E(X)$
- Stichprobenvarianz
  $$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \text{ für Varianz}$$

} erwartungstreu

- prüfen: erwartungstreu wenn bei E(Schätzer) gewünschter Parameter rauskommt

Konsistenz

: je mehr X_i ich habe, desto genauer muss Schätzer sein

⟹ Kriterium der Nichtkonsistenz: verwendet nicht alle geg. X_i

⟹ prüfen sonst: für alle Komponenten ihre Grenzwerte einsetzen. wenn dann gewünschter Parameter rauskommt: konsistent

⟶ Stetigkeitssatz: stetige Funktion auf konsistenten Schätzer anwenden ↳ bleibt konsistent

Konfidenzintervalle

⟹ Transformation von KI: $[G_u, G_o]$ KI zu θ, h streng monotone Funktion

dann ist $[h(G_u), h(G_o)]$ wachsend, fallend $[h(G_o), h(G_u)]$ ein KI zu $T = h(\theta)$

⟹ Genauigkeit = halbe Länge eines KI

wie viele n brauche ich für Genauigkeit Δ? mit Schätzer $\hat{\pi}$

1) mit Vorerhebung / Schätzwert ($\hat{\pi}_*$)

$$n \geq \hat{\pi}_* \cdot (1 - \hat{\pi}_*) \cdot (z_{1-\alpha/2} / \Delta)^2$$

2) ohne Vorinfo „worst case" evtl zu viel

$$n \geq \frac{1}{4} \cdot (z_{1-\alpha/2} / \Delta)^2$$

Tests:

= formale Entscheidungsregel (H_0 / H_1) basierend auf einer ZV = Teststatistik / Prüfgröße

- Statistisches Modell, Verteilungsannahme
- Nullhypothese H_0 + Alternative H_1 – schließen sich aus, decken $\mathbb{R}$ ab
- zweiseitiger Test: $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$ — θ_0 = konkrete Zahl!
- einseitiger Test:
  - a) $H_0: \theta \leq \theta_0$ $H_1: \theta > \theta_0$
  - b) $H_0: \theta \geq \theta_0$ $H_1: \theta < \theta_0$

  zu zeigende Aussage in H_1 formulieren!
- Fehler:

wahr \ entschieden für	H_0	H_1
H_0	✓	Fehler 1. Art $P = \alpha$ unabh. von n!
H_1	Fehler 2. Art $P = ?$ ↳ kleiner bei n groß	✓

„Signifikanztest zum Niveau α“ $P(\text{Fehler 1. Art}) \leq \alpha$

→ wenn für H_1 entschieden → Fehlerquote ~~max.~~ α „statistisch bewiesen“ / „signifikant“

↳ unsicher, Tendenz

- Gauß-Tests: = z-Test

$X \sim N(\mu, \sigma^2)$, σ^2 bekannt, herausfinden: μ

Hypothesen:
- a) $H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$
- b) $H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$
- c) $H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$

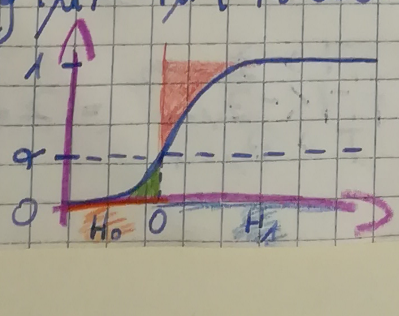
Teststatistik: $T = \dfrac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}}$

kritische Werte: H_1 wahr wenn:
- a) $T < -z_{1-\frac{\alpha}{2}}$ / $T > z_{1-\frac{\alpha}{2}}$
- b) $T < z_{\alpha}$
- c) $T > z_{1-\alpha}$

Gütefunktion

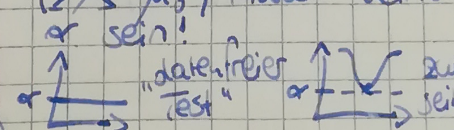
Bsp.: $H_0: \mu \leq 0$ $H_1: \mu > 0$

$g(\mu) = P_\mu(\text{"}H_0\text{ ablehnen"})$



- Fehler 1. Art (max. α)
- Fehler 2. Art (1 − Funktion = β)

· Merkmale:
- i.d.R. stetig
- Werte zw. 0 & 1 (weil Wahrsch.)
- je besser, desto steiler
- alle Werte, die H_0 annehmen ($\geq$ / $\leq$ μ_0) müssen kleiner α sein!



p-Werte

= 1. Wert, zu welchem α H_1 angenommen wird (1. = kleinster)
↳ aber 1. α festlegen
2. p-Wert ermitteln

· man entscheidet zugunsten H_0 wenn μ_0 im KI liegt beim 2-seitigen Problem

· p-Wert < α: H_1
p-Wert > α: H_0

↳ für alle Test gültig

p-Werte für Gauß-Test → bei jedem Test anders

a) $H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$ p-Wert = $2(1-\Phi(|T|))$
b) $H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$ p-Wert = $\Phi(T)$
c) $H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$ p-Wert = $1-\Phi(T)$

T = Wert der Teststatistik/(Prüfgröße)

Einstichproben t-Test

$X_i \sim N(\mu, \sigma^2)$ Normalverteilung vorausgesetzt, aber unbekannte Varianz

$$T = \frac{\bar{X}_n - \mu_0}{\sqrt{S_n^2 / n}} = \sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n}$$

→ falls $n > 30$ z (Φ) nehmen!

H_0 ablehnen falls:
a) $|T| > t_{n-1,1-\alpha/2}$
b) $T < -t_{n-1,1-\alpha}$ $(= t_{n-1,\alpha})$
c) $T > t_{n-1,1-\alpha}$

Approximativer Gauß-Test

$n \geq 30$, beliebige Verteilung, unbekannte Varianz

T = analog t-Test, H_0 ablehnen falls:

a) $|T| > z_{1-\alpha/2}$
b) $T < -z_{1-\alpha}$
c) $T > z_{1-\alpha}$

nur limes von Fehler 1. Art $\leq \alpha$! wenn $n \to \infty$.

Approximativer Binomialtest

$X_i \sim B(1, \pi)$, $n \geq 30$, $0 < \pi_0 < 1$ (Wahrsch.)

$$T = \frac{X/n - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}} = \frac{X - n \cdot \pi_0}{\sqrt{n \cdot \pi(1-\pi_0)}}$$

H_0 ablehnen falls:
a) $|T| > z_{1-\alpha/2}$
b) $T < -z_{1-\alpha}$ $(= z_\alpha)$
c) $T > z_{1-\alpha}$

Einstichprobentests

Zweistichprobentests

Gauß-Test

normalverteilt, bekannte Varianzen

δ_0 (Delta) $= \mu_x - \mu_y \rightarrow$ wie soll es sein?

Nachweis, dass kein Unterschied: 0
Nachweis, dass μ_y um 1% kleiner als μ_x: $-0{,}01$
etc.
! nicht Erwartungswerte berechnen !

a) $H_0: \mu_x - \mu_y = \delta_0$ $H_1: \mu_x - \mu_y \neq \delta_0$

b) $H_0: \mu_x - \mu_y \geq \delta_0$ $H_1: \mu_x - \mu_y < \delta_0$

c) $H_0: \mu_x - \mu_y \leq \delta_0$ $H_1: \mu_x - \mu_y > \delta_0$

$$T = \frac{\overline{X}_n - \overline{Y}_m - \delta_0}{\sqrt{\sigma_x/n + \sigma_y/m}}$$

Ablehnen H_0 als falls:

a) $|T| > z_{1-\alpha/2}$

b) $T < -z_{1-\alpha}$ $(= z_\alpha)$

c) $T > z_{1-\alpha}$

2-Stichproben-t-Test

normalverteilt, unbekannte, aber gleiche Varianzen

$$T = \frac{\overline{X}_n - \overline{Y}_m - \delta_0}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \cdot S_p^2}}$$

$$S_p^2 = \frac{1}{n+m-2}\left(\sum_{i=1}^{n}(x_i - \overline{x}_n)^2 + \sum_{i=1}^{m}(y_i - \overline{y}_m)^2\right) = \frac{1}{n+m-2}\left(n \cdot \tilde{S}_x^2 + m \cdot \tilde{S}_y^2\right)$$

H_0 ablehnen falls:

a) $|T| > t_{n+m-2,\,1-\alpha/2}$

b) $T < -t_{n+m-2,\,1-\alpha}$

c) $T > t_{n+m-2,\,1-\alpha}$

approximativer Gauß-Test

beliebige Verteilung, unbekannte Varianzen

$$T = \frac{\overline{X}_n - \overline{Y}_m - \delta_0}{\sqrt{\tilde{S}_x^2/n + \tilde{S}_y^2/m}}$$

H_0 ablehnen falls: siehe Gauß-Test!

Kritische Werte

für alle Tests

a) $|T| > z_{1-\alpha/2}$
b) $T < -z_{1-\alpha}$
c) $T > z_{1-\alpha}$

Gauß-Test
approx. Gauß-Test
approx. Binomialtest
Zweistichproben Gauß-Test
approx. Zweistichproben Gauß-Test
approx. Zweistichproben Binomialtest
sonst. Korrelationstest

a) $|T| > t_{n-1, 1-\alpha/2}$
b) $T < -t_{n-1, 1-\alpha}$
c) $T > t_{n-1, 1-\alpha}$

Einstichproben t-Test
t-Test verbundene Stichproben

a) $|T| > t_{n+m-2, 1-\alpha/2}$
b) $T < -t_{n+m-2, 1-\alpha}$
c) $T > t_{n+m-2, 1-\alpha}$

Zweistichproben t-Test

- $T > \chi^2_{(k-1)(m-1), 1-\alpha}$

χ^2-Unabhängigkeitstest

a) $|T| > t_{n-2, 1-\alpha/2}$
b) $T < -t_{n-2, 1-\alpha}$
c) $T > t_{n-2, 1-\alpha}$

Korrelationstest
ob ρ_{XY} größer/kleiner 0

Tests für α & β