Wahrscheinlichkeiten und Statistik

Fabio Oesch

11. Dezember 2012

Inhaltsverzeichnis

1 Themen

1. beliebig oft wiederholbar
2. Resultat ist zufällig

Bsp: Lotto, Münze werfen, Würfeln

1.1 Ergebnisraum

$\Omega =$ Menge aller möglichen Ausgänge des Experiemnts (im Skript mit S)
Bsp: Ω von Lotto: $\Omega = \{1, \dots, 45\}$

1.2 Ereignis $E \subseteq \Omega$

Ereignis $E =$\{Augenzahl ist gerade\}
spezielle Ereignisse:

- $E = \Omega$ sicheres Ereignis
- $E = \emptyset$ unmögliches Ereignis

E Ereignis: $E^C = \overline{E} = \Omega \backslash E$

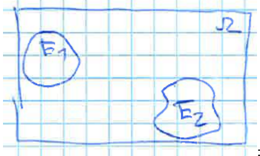
1.3 Wahrscheinlichkeitsfunktion

$\mathbb{P} : \mathcal{P}(\Omega) \to [0, 1]$
2 wichtige Eigenschaften: $\mathcal{P}(\Omega) = 1$, $\mathcal{P}(\emptyset) = 0$, $0 \leq \mathcal{P}(E) \leq 1$
Bsp: $\Omega = \{$Kopf, Zahl, Kante$\}$

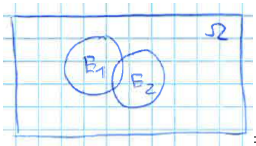
$$\left.\begin{array}{l} \mathcal{P}(\{\text{Kopf}\} = \frac{2}{3} \\ \mathcal{P}(\{\text{Kante}\} = 0 \\ \mathcal{P}(\{\text{Zahl}\} = \frac{1}{3} \end{array}\right\} \Rightarrow \mathcal{P}(E) = 1,\ E = \{\text{Kopf, Zahl}\}$$

1.4 Subaditivität

$E_1,\ E_2 \subseteq \Omega,\ E_1 \cap E_2 = \emptyset \Rightarrow \mathcal{P}(E_1 \cup E_2) = \mathcal{P}(E_1) + P(E_2)$



$\Rightarrow \mathcal{P}(E_1 \cup E_2) = \mathcal{P}(E_1) + \mathcal{P}(E_2),\ E = E_1 \cup E_2$



$\Rightarrow \mathcal{P}(E) = \mathcal{P}(E_1) + \mathcal{P}(E_2) - \mathcal{P}(E_1 \cap E_2),\ E = E_1 \cup E_2$
$\mathcal{P}(E_1 \cup E_2 \cup \dots \cup E_n) \leq \mathcal{P}(E_1) + \dots + \mathcal{P}(E_n)$
Bsp: Lotto mit Matryoshka

$$\mathcal{P}(\{w\}) = \frac{1}{45},\ E_1 = \{1\},\ E_2 = \{1, 2\},\ \dots,\ E_{45} = \{1, \dots, 45\} = \Omega \Rightarrow \mathcal{P}\overbrace{(E_1 \cup \dots \cup E_{45})}^{\Omega} = 1$$
$$\mathcal{P}(E_1 \cup E_2 \cup \dots \cup E_{45} \leq \mathcal{P}(E_1) + \dots + \mathcal{P}(E_{45}) = \frac{1}{45} + \frac{2}{45} + \dots + \frac{45}{45} = \frac{\frac{45 \cdot 46}{2}}{45} = \frac{46}{2} = 23$$

1.5 Wahrscheinlichkeitsraum (Vorläufig)

$W = (\Omega,\ \mathcal{P}(\Omega),\ \mathbb{P})$, $\Omega =$ Ergebnisraum, $\mathcal{P}(\Omega)$ =alle Ausgänge des Experiments. alle E's, $\mathbb{P}$ = Wahrscheinlichkeitsfunktion

1.6 Theoretische Wahrscheinlichkeitsfunktion

$\mathcal{P}(\{\text{Zahl}\}) = \mathcal{P}(\{\text{Kopf}\}= \frac{1}{2}$ (Definiere die Wahrscheinlichkeit synthetisch)
$\mathcal{A}(E) = \frac{\text{wie häufig tritt } E \text{ ein bei N-facher Wiederholung}}{N}$ (Empirische Wahrscheinlichkeit)

1.7 Laplace-Experimente

$\triangleq$ Fairen Spielen, Die Wahrscheinlichkeiten sind gleichverteilt
$|\Omega| = n$ endlicher Wahrscheinlichkeitraum.
Jedes Elementarergebnis ist gleich wahrscheinlich ($|E| = 1$). **Bsp:** Würfel: $\{1\} = \{1, 2, 3, 4, 5, 6\}$. kein Elementarergebnis: $\{3, 4\}$ $\omega \in \Omega : \mathcal{P}(\{\omega\}) = \frac{1}{n} = \frac{1}{|\Omega|}$, $|A|$= Anzahl Elemente in A

1.8 Mehrstufige Zufallsexperiment

Zufallsexperiment Z, das mehrfache hintereinander angeführt wird.
Bsp: mehrmals Würfeln: Wie gross ist die W'keit $2\times$ hintereinander 6 zu würfeln: $\mathcal{P}(2\times \text{ 6 Würfeln}) = \frac{1}{36}$
Produktregel: $\mathcal{P}(E_1 \text{ und } E_2) = \mathcal{P}(E_1) \cdot \mathcal{P}(E_2)$
Möglichkeiten: Ω_1 hat n_1 viele Ausgänge ($|\Omega_1| = n_1$), Ω_2 hat n_2 viele Ausgänge ($|\Omega_2| = n_2$) also $n_1 \cdot n_2$

1.9 Urnenmodel

Unterscheidung nach „Zurücklegen“ oder „nicht zurücklegen“ und „geordnet“ oder „keine Reihenfolge“

	zurücklegen	nicht zurücklegen
geordnet	n^k	$n!$ oder $\frac{n!}{(n-k)!}$
ungeordnet		$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

$p = \frac{\text{Anzahl der gnstigen Flle}}{\text{Anzahl der mglichen Flle}} = \frac{g}{m}$
Bsp: Klasse aus 10 Mädchen und 14 Knaben. Wähle 5 Personen aus.
a) W'keit, dass alle Mädchen sind?
Antwort: $P(\{\text{5 Mädchen}\}) = \frac{|\{\text{5 Mädchen}\}|}{|\Omega|}$, $|\Omega| = \binom{24}{5}$, $|\{\text{5 Mädchen}\}| = \binom{10}{5}$
$\Rightarrow P(\{\text{5 Mädchen}\}) = \frac{\binom{10}{5}}{\binom{24}{5}}$
b) W'keit alles Knaben: $P(\{\text{5 Knaben}\}) = \frac{\binom{14}{5}}{\binom{24}{5}}$
c) W'keit, dass in der 5-er Gruppe, sowohl Mädchen, als auch Knaben vorkommen.
Gegenw'keit von a) + b), also $\overline{E} = \{\text{nur Mädchen oder nur Knaben}\} \Rightarrow P(\overline{E}) = P(\{\text{nur Mädchen}\}) + P(\{\text{nur Knaben}\}) = \frac{\binom{10}{5}+\binom{14}{5}}{\binom{24}{5}} \Rightarrow P(E) = 1 - P(\overline{E})$
Gegenw'keit benutzen: $P(E)$, $\frac{\Omega}{E} = \overline{E}$, $1 - P(\overline{E}) = P(E)$
Bsp: 8x Münze werfen
Wie gross ist die W'keit, das Zahl & Kopf gleichhäufig vorkommen.
$|E| = \binom{8}{4} \cdot \binom{4}{4}$

1.10 Repetition

$$\left.\begin{array}{l} X \text{ Zufallsvariable: Anzahl bei 1x würfeln} \\ Y \text{ Zufallsvariable: Anzahl bei 1x würfeln} \end{array}\right\} X, Y \text{ Gleichverteilt (uniform) uniform}$$

$Z := X + Y$ Zufallsvariable: $F(z) = \sum_{X_i \leq Z} p_i$, $p_i = P(Z = x_i)$

1.11 Erwartungswert:

Theoretischer Pendant zum Mittelwert.
Bsp: x_1, x_2, x_3, x_4, $h = |\{x_1, x_2, x_3, x_4\}| \to \overline{x} = \frac{\sum x_i}{h}$
$\mathbb{E}X = \mu = \sum_{\text{alle } X_i} x_i \cdot \mathbb{P}(X = x_i)$ keine Zufallsvariable
Bsp: X sein die Augenzahl von 1x würfeln. $\mathbb{E}X = \sum_{i=1}^{6} i \cdot \frac{1}{6} = 3.5$
$Y = X - \mathbb{E}(X)$ ist eine Zufallsvariable, $\mathbb{E}Y = \mathbb{E}(X - \mathbb{E}X) = \sum_{i=1}^{n}(x_i p_i - p_i \mathbb{E}(X)) = \sum_{i=1}^{n} x_i p_i -$

$\sum_{i=1}^{n} p_i \mathbb{E}(X) = \mathbb{E}(X) - \mathbb{E}(X) \sum_{i=1}^{n} p_i = \mathbb{E}(X) - \mathbb{E}(X) = 0$
Bsp: Würfel $\mathbb{E}(X) = 3.5$, $E(X - \mathbb{E}(X)) = 0$

$X - 3.5$	$1 - 3.5$	$2 - 3.5$	$3 - 3.5$	$4 - 3.5$	$5 - 3.5$	$6 - 3.5$
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

1.12 Streuung, Varianz

$\mathrm{Var}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) \Rightarrow \mathbb{E}(X-\mu)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$

1.13 Binomialverteilung

Zufallsexperiment mit 2 Ausgängen: Erfolg, Misserfolg
$P(X = \text{Erfolg}) = p \in [0,1]$, $P(X = \text{Misserfolg}) = 1 - p = q$
Zufallsvariable X = Anzahl Erfolge bei n-facher Wiederholung des Experiments
Wahrscheinlichkeitsfunktion von X aus. $P(X = x) =$

x_i	0	1	k	n
p_i	$(1-p)^n$	$\binom{n}{1}p(1-p)^{n-1}$	$\binom{n}{k}p^k(1-p)^{n-k}$	$\binom{n}{n}p^n$

Bsp:
Spieler A: M.D. 40% Erfolgsw'keit, Spieler B: K.G. 60% Erfolgsw'keit.
Sie spielen 3x gegeneinander. E: W'keit dass A häufiger als B gewinnt. Also muss A 2- oder 3-Mal gewinnen. $P(X = 2) + P(X = 3) \Rightarrow \binom{3}{2}0.4^2 \cdot 0.6 + \binom{3}{3}0.4^3 + 0.6^0 = 0.352 \Rightarrow 35.2\%$ W'keit gewinnt A
Erfolgsw'keit von p: $\mathbb{E}X$ bei n spielen. Mit Trick: $\mathbb{E}X = n \cdot p$
$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$
$f(t) = (q + pt)^n = \sum_{k=0}^{n} \binom{n}{k} q^{n-k}(pt)^k$

1.13.1 Binomialverteilung:

Abkürzung: $X \sim \mathrm{Bin}(n,p)$, n: Anzahl Experimente, p: Erfolgsw'keit
$P = (X = k) = \binom{n}{k}p^k(1-p)^{n-k}$, $\mathbb{E}(X) = np$, $\mathrm{Var}(X) = \sigma^2 = np(1-p)$. σ: Std'abweichung

1.13.2 Testen einer Hypothese:

Vermutung: Hühner können zw. $\circ$ und $\triangle$ Futter entscheiden.
$20 \times \circ$, $20 \times \triangle \Rightarrow \circ =$ Erfolg, $\triangle =$ Misserfolg
Zufallsvariable X zählt die Anzahl Erfolge $\Rightarrow$ Binomiales Experiment d.h. $X \sim \mathrm{Bin}(20, p)$.
Führen das Experiment durch: $15 \times \circ$ und $5 \times \triangle$, experimentelle W'keit für Erfolg: $p = \frac{15}{20} = \frac{3}{4}$
Hypothese formulieren:
H_0: Nullhypothese: es gibt keinen Unterschied $\rightarrow$ Huhn kann nicht unterscheiden zw. $\circ$ & $\triangle$, $p = q = \frac{1}{2}$
H_1: Alternativhypothese: $p \geq q$ ($p \leq q$). d.h. es gibt einen Unterschied beim Fressverhalten.
Ziel: Entscheiden ob H_0 anzunehmen ist, oder sie zugunsten von H_1 verwerfen.
Berechnung: Berechne W'keit unter $H_0(p = q = \frac{1}{2})$, dass wir einen Ausgang mit $15\times$ Erfolg und $5\times$ Misserfolg
$P(15 \leq X \leq 20) = \sum_{k=15}^{20} \binom{20}{k} p^k (1-p)^{20-k} \overset{\text{unter } H_0!}{=} \sum_{k=15}^{20} \binom{20}{k} \frac{1}{2}^k \cdot \frac{1}{2}^{20-k} \approx 0.021 = 2.1\%$, Signifikanz-Niveau α, $\alpha = 0.1$ (10%) $\Rightarrow$ Falls $P(15 \leq X \leq 20 | H_0) \leq \alpha \Rightarrow$ dann verwerfen H_0, ansonsten nehmen wir H_0 an.
• 2 Möglichkeiten: falls unter der Nullhypothese

1. $P(15 \leq X \leq 20) > \alpha$, dann nehmen wir die Nullhypothese an es spricht nichts gegen H_0 auf Signifikanzniveau α
2. $P(15 \leq X \leq 20) \leq \alpha$, dann verwerfen wir H_0 zugunsten von H_1

• (2) Fehler 1. Art; verwerfen von H_0, obwohl H_0, obwohl H_0 korrekt wäre $\rightarrow$ Irrtumsw'keit $P(15 \leq X \leq 20)$.
• (1) Fehler 2. Art; verwerfen H_1, obwohl H_1 korrekt ist $\rightarrow$ Irrtumsw'keit β (Power)

1.14 Poissonverteilung

1. Gleichverteilung (fairer Wrfel)
2. Binomialverteilung
3. Poissonverteilung

Idee: p soll sehr klein sein. n soll sehr gross sein.
Bsp: X sei binomialverteilt und Parametern n, p. $\Rightarrow \mathbb{E}(X) = n \cdot p$.
$\lim_{n\to\inf} n \cdot p_n = \lambda \in \mathbb{R}, \Rightarrow p = \frac{\lambda}{n}$
X binomialverteilt: $P(X = k) = \binom{n}{k} \cdot p^k(1-p)^{n-k}, \overset{?}{\Rightarrow} \lim_{n\to\inf} P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$
Verteilung mit W'keitsfunktion $\frac{\lambda^k}{k!} \cdot e^{-\lambda}$ heisst Poissonverteilung.
Erwartungswert von $X\tilde{P}oi(\lambda)$ (X ist Poisson-verteilung mit Parameter λ)
$\mathbb{E}(X) = \sum_{x=0}^{\infty} x \cdot P(X = x) = \lambda$

- Fr sehr kleine W'keiten. Mit bekanntem „Mittelwert" (Erwartungswert) λ. Mit quasi unendlich (unbekannter) Anzahl gleicher Experimente.
- $P(X = x) = \frac{\lambda^x e^{-lambda}}{x!}$
- $\mathbb{E}(X) = \lambda$
- $\mathrm{Var}(X) = \lambda$
- $X\tilde{P}oi(\lambda)$, $Y\tilde{P}oi(\mu)$, $Z = X + Y : Z\tilde{P}oi(\lambda + \mu)$

Bsp: Smartphonehersteller, Fehlerquote von 1, 5.000 Smartphones
Wie gross ist die W'keit, dass mind. 2 defekt sind.
Poissonapproximation: $\lambda = 0.001 \cdot 5000 = 5$
$P(X \geq 2) = 1 - (P(X = 1) + P(X = 0)) = 1 - (\frac{\lambda^1 e^{-\lambda}}{1!} + \frac{\lambda^0 e^{-\lambda}}{0!}) = 1 - (\frac{5^1 e^{-5}}{1!} + \frac{5^0 e^{-5}}{0!}) = 1 - 6e^{-5} \approx 0.96$

1.15 Statistische Tests

1.15.1 Prinzip des statistischen Tests

1. Nullhypothese H_0 formulieren.
   Bsp: $\left.\begin{array}{r} H_0 : p = \frac{1}{6} \text{ bei einem Wrfel} \\ H_1 : p > \frac{1}{6} \text{ Alternativhypoth.} \end{array}\right\} H_1 \cap H_0 = \emptyset$
2. Signifikanzniveau $\alpha \in (0, 1)$
3. Stichprobe sammeln
4. Entscheid fllen: Berechne W'keit unter H_0, das wir einen Ausgang haben, wie die Stichprobe
   Ist sie grsser als $\alpha \Rightarrow H_0$ annehmen
   Ist sie kleiner als $\alpha \Rightarrow H_0$ zugunsten von H_1 verwerfen

1.15.2 Beispiel

12'000 mal Wrfeln, 2'107 mal Sechs, $\alpha = 10\%$

- $H_0 : p = \frac{1}{6}$ Nullhypothese
- $H_1 : p > \frac{1}{6}$ Alternativhypothese

$\mathbb{E}(X) = np = 12000 \cdot \frac{1}{6} = 2000$
Wir berechnen die Wahrscheinlichkeit von $P(2107 \leq X)$. Wir bentigen Grenzwertsatz von de Moivre und Laplace. $\mu = np = 12000 \cdot \frac{1}{6} = 2000$ und $\sigma^2 = np(1-p) = 12000 \cdot \frac{1}{6} \cdot \frac{5}{6} = 1666\frac{2}{3}$. Die Approximation ist erlaubt, da die Faustregel $np(1-p) = 1666\frac{2}{3} > 9$ erflt ist.
Nun erhalten wir mit der Tafel:
$P(2107 \leq X) \approx 1 - \Phi(\frac{2107-2000}{\sqrt{1666\frac{2}{3}}}, 0, 1) = 1 - \Phi(2.621, 0, 1) = 1 - 0.9956 = 0.0044$
Da $\alpha < 0.0044$ ist verwerfen wir H_0

1.15.3 Erwartungswert und Varianz

Auf dem Intervall $I =]-\infty, \infty[$
Erwartungswert $\mu = \mathbb{E}(X) = \int_{x_1}^{x_2} xf(x)dx$, Varianz $\sigma^2 = \int_{x_1}^{x_2} x^2 f(x)dx - \mu^2$

1.16 t-Verteilung

$X_1, \ldots, X_n \sim N(\mu, \sigma^2)$
$T_{n-1} := \frac{\frac{1}{n}\sum_{i=1}^n X_i - \mu}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n (X_i-\mu)^2}}\sqrt{n} \in \mathbb{R}^{\Omega^n}$ heisst t-verteilt mit $n-1$ Freiheitsgraden
$t_{n-1} := \frac{\frac{1}{n}\sum_{i=1}^n x_i - \mu}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n (x_i-\bar{x})^2}}\sqrt{n} \in \mathbb{R}$
$f_{n-1}(t) = c_{n-1}(1+\frac{t^2}{n})^{-\frac{n+1}{2}}$ fr $\lim_{n\to\infty} \to \varphi(t,0,1)$
$c_{n-1} = \frac{\Gamma(\frac{n}{2})}{\sqrt{\pi(n-1)}\Gamma(\frac{n-1}{2})}$, $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$. Gammafunktion.
$\Gamma(x+1) = x\Gamma(x)$; $\Gamma(1) = 1$; $\Gamma(5) = \Gamma(4+1) = 4\Gamma(4) = 4\Gamma(3+1) = 4\cdot 3\cdot\Gamma(3) = \cdots = 5!$

1.17 Parametertests

- 1-Stichprobentest
  $H_0 : \mu = \bar{x}$ $X \sim W(\mu_1, \sigma_1^2)$ $H_1 : \mu \neq \bar{x}$ $Y \sim W(\mu_2, \sigma_2^2)$
- 2-Stichprobentest
  $H_0 : \mu_1 \neq \mu_2$

Grundvoraussage:

- Verteilungsfamilie bekannt
  (d.h. $W(\cdot,\cdot)$, t-verteilt, Weibull etz)
- Testen ob 1-Stichpr.fall $H_0 : \vartheta = \hat{\vartheta}_n$ (ϑ: fester Wert, $\hat{vartheta}$: empirisch)
  2-sTichpr.fall $H_0 : \hat{\vartheta}_n = \hat{\vartheta}_m$

1. Verteilung von T_n (Teststatistik) unter H_0 bekannt $\Leftrightarrow$: exakter Test.
2. Verteilung von T_n unter H_0 unbekannt.

Nicht parametrische Tests (Verteilungsfreie Tests):

- 1./2. Stichprobentests existieren
- keine Verteilungsparameter

1. Stichprobentest $H_0 : \hat{F_n}(x) = F_0(x)$
2. Stichprobentest $H_0 : \hat{F_n}(x) = \hat{G_m}(x)$
$\bar{F_n}(x) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}\{x_i \leq x\}$

1.18 Konfidenz/Vertrauensintervall

$X \sim N(\mu, \sigma^2)$
Experiment n-mail durchfhren. $\Rightarrow X = \{x_1, \ldots, x_n\}$
Theoretisch $\mathbb{E}(X) = \mu$
Empirisch: $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$ als Schtzung von μ
Wie gut ist die Schtzung? Vorgabe: $\gamma \in [0,1]$
$[\mu - \triangle x, \mu + \triangle x] \ni \bar{x}$ Finde $\triangle x$, so dass $\mathbb{P}(\bar{x} \in [\mu - \triangle x, \mu + \triangle x] = \gamma$.
$\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$. $s^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - x)^2$. $t_{n-1} = \frac{\bar{x}-\mu}{s}$ (t-verteilt (exakt) fr ZV.).
$\mathbb{P}(\bar{x} \in [\mu - \triangle x, \mu + \triangle x]) = \gamma = 1 - \alpha$
Beispiel: $n = 10$ Stichproben $X = \{x_1, \ldots, x_n\}$
$\bar{x} = 5$. $s = 0.2$ Vertrauensintervall =? bei Vertraunsw'keit von $\gamma = 0.95 \Rightarrow \alpha = 0.05$; $1 - \frac{\alpha}{2} = 1 - 0.025 =$

0.975
$\Rightarrow \triangle x = \frac{t_{10-1,0.975}}{\sqrt{10}} \cdot 0.2$ Tabelle betrachten Seite 140 ergibt mit Freiheitsgrad $= n-1 = 9$. $t_{9,0.975} = 2.262$.
$\Rightarrow \triangle x = \frac{2.262}{\sqrt{10}} \cdot 0.2 \approx 0.14$
Vertrauensintervall $[4.86, 5.14]$. mit W'keit von 95% liegt μ in $[4.86, 5.14] = [5- \triangle x, 5+ \triangle x]$
Allgemein
Feste Vertrauensw'keit $\gamma = 0.95$
s konstant
Stichproben n: $\triangle x = \frac{t_{N-1,0.975}}{\sqrt{N}} s(n \to \infty) = 0$
$t_{n,0.975} \leq t_{n-1,0.975} \leq 2.262$ fr $n \geq 10$
Beispiel:

$$t = \frac{\overbrace{\bar{x} - \mu}^{\triangle\mu}}{s} \sqrt{s}$$

Festes γ, und $\triangle \mu$. Wie gross n whlen?
$\frac{t \cdot s}{\bar{x} - \mu} = \sqrt{n} \overset{2}{\Rightarrow} \frac{t^2 \cdot s^2}{(\Delta x)^2} \leq n$
• Vorgehen: - Vertrauensintervallgrsse $\triangle \mu$, - Vertrauensw'keit γ ($\rightarrow$ fliesst in t ein)
$t = t_{n-1,1-\frac{1-\gamma}{2}}$ Quantifunktion der t-Verteilung.
• Abschtzung: $t \approx 2$ (oder $t = 3$). Daumenregel: $\frac{4 \cdot s^2}{(\Delta\mu)^2} \geq N$ (s^2 ist geschtzt)

1.18.1 2-Stichproben t-Test

Annahme: • Normalverteilung, • 2 Gruppen
$H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$
2 Flle: a) unbekannte, aber gleiche Varianz d.h. $\sigma_1^2 = \sigma_2^2 = \sigma_2$ (Homoskedastisch) exakter Test
b) unbekannte, evtl. ungleiche Varianzen, $\sigma_1^2 \neq \sigma_2^2$ (Heteroskedastisch) approximation

1.18.2 Homoskedastischer Fall

Zwei Stichproben: $X = \{x_1, \ldots, x_n\}$ $W(\mu, \sigma^2)$; $Y = \{y_1, \ldots, y_n\}$ $W(\mu_2, \sigma^2)$
μ_1, μ_2, σ^2 sind unbekannt.
Testen, ob $\mu_1 = \mu_2$ auf Signifikanzniveau α (2-seitiger Test, d.h. $H_0 : \mu_1 = \mu_2$ gegen $H_1 : \mu_1 \neq \mu_2$)
Testgrsse: $t = \frac{\bar{x} - \bar{y}}{s} \sqrt{\frac{n \cdot m}{n+m}}$. t-verteilt mit $n + m - 2$ Freiheitsgraden
$\bar{s_1} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$, $\bar{m - 1} = \frac{1}{m} \sum_{i=1}^{m} (y_i - \bar{y})^2$
$s = \frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2 + \sum_{i=1}^{m} (y_i - \bar{y})^2}{n+m-2}$
Signifikanzniveau α: Ist $|t| < t_{n+m-2,1-\frac{\alpha}{2}} \Rightarrow H_0$ annehmen. Ist $|t| \geq t_{n+m-2,1-\frac{\alpha}{2}} \Rightarrow H_0$ ablehnen.

2 Beispiele:

Aufgabe 3.5.8. Auf wie viele Arten können die Buchstaben des Wortes Pfeffer permutiert werden? $\frac{7!}{3! \cdot 2!} = 420$
Aufgabe 3.5.10. Eine Klasse hat 15 Fussballspieler, einer davon heisst Klaus. Auf wie viele Arten kann eine Mannschaft von 11 Spielern a. Mit Klaus, b. ohne Klaus zusammengestellt werden. a. $\binom{14}{10}$, b. $\binom{14}{11}$
Aufgabe 3.5.14. Wie viele Möglichkeiten gibt es, die 36 Jasskarten auf vier Spieler A, B, C, D zu verteilen? $\binom{36}{9} \cdot \binom{27}{9} \cdot \binom{18}{9} \cdot \binom{9}{9}$
Aufgabe 4.1.3 von 10 Nssen sind 3 verdorben. Wahrscheinlichkeit, dass 2 gute Nsse genommen werden, $m = \binom{10}{2}, g = \binom{7}{2} \Rightarrow p = \frac{\binom{7}{2}}{\binom{10}{2}} \frac{7}{15}$
Aufgabe 4.1.12 10 Lose 2 Gewinnlose mit 5x herausziehen genau 1 Gewinnlos $\Rightarrow m = 8, n = 2, k = 5, s = 1 \Rightarrow \frac{\binom{8}{4}\binom{2}{1}}{\binom{10}{5}}$
Aufgabe 4.1.13 Es liegen $m+n$ Lose vor, unter denen n Gewinnlose sind. Es werden k Lose auf einmal gezogen. Bestimmen Sie die W'keit dafür, dass sich unter den k Losen genau s Gewinnlose befinden.
Es gibt $m = \binom{m+n}{k}$ mögliche Ausfälle und genau $g = \binom{m}{k-s}\binom{n}{s}$ Möglichkeiten für s Gewinnlose. Dabei zählt der Faktor $\binom{m}{k-s}$ die Möglickeiten, $k-s$ Nieten zu haben und der Faktor $\binom{n}{s}$ zählt die Möglichkeiten,

s Treffer zu haben. Damit folgt $p(s) = \frac{\binom{m}{k-s}\binom{n}{s}}{\binom{m+n}{k}}$.

Bsp: Wrfel fr sechs $p = \frac{1}{6}, p > \frac{1}{6}$ a) 2x 6 in 3 Wrfen b) 3x 6 in 5 Wrfen
$H_0 : p = \frac{1}{6}$ mit $\alpha = 5\%$ also gilt: a) $P(2 \leq X \leq 3) = \binom{3}{2}(\frac{1}{6})^2(\frac{5}{6})^1 + \binom{3}{3}(\frac{1}{6})^3(\frac{5}{6})^0 = 0.074 > \alpha \Rightarrow H_0$ angenommen; b) $P(3 \leq X \leq 5) = \binom{5}{3}(\frac{1}{6})^3(\frac{5}{6})^2 + \binom{5}{4}(\frac{1}{6})^4(\frac{5}{6})^1 + \binom{5}{5}(\frac{1}{6})^5(\frac{5}{6})^0 = 0.035 < \alpha \Rightarrow H_0$ nicht angenommen;