Schliessende Statistik

Parametrische vs. Nichtparametrische Modelle

- **Parametrische Modelle**
  - Annahme über zugrundeliegende Verteilung (Binomial-/Normalverteilung, etc.) wird vorausgesetzt.
  - Durch diese Annahme ist das Modell genauer
  - Suchen mindestens einen Parameter (μ, p, etc.)
  - Schätzung der unbekannten Parameter liefert ein vollständig spezifiziertes Modell
- **Nichtparametrische Modelle**
  - Machen **keine** Annahme über zugrundeliegende Verteilung.
  - Sind an Kenngrösse interessiert (E(X), Var(X), Median, Quantile, etc.).
  - Modell ist universeller anwendbar, aber dafür mit einer verminderten Macht
  - Schätzen der Kenngrösse liefert kein komplettes Modell

Packages

- boot
- BSDA
- car
- coin
- fitdistrplus

Modellwahl und Überprüfung der Modellannahmen

- **Verteilungsannahme → qqPlot(); qqplot(); descdist()**
  - Wichtig bei kleinen Stichproben, Asymmetrien, Lang-schwänzigkeiten, sonst ZGWS → Verteilung nicht wichtig
  - Falls Verteilung unbekannt → Cullen-Frey-Diagramm
  - qqplot → Vergleich, ob zwei Stichproben aus selben Vert.
  - qqPlot → Vergleich 1 Stichprobe mit beliebigen Standart-Verteilung oder Vert. mit beliebigen Werten (Steigung = 1!)
- **Unabhängigkeit → scatterplot(); acf(); cor()**
  - Abhängigkeit festzustellen kann schwierig sein
  - Kann davon abhängig sein, wie Daten sortiert sind
  - Untersuchung mit scatterplot / Auto-Korrelationsfunktion
  - ACF: Wenn keine Korrelationen, sollten Werte sehr nahe bei null liegen, idealerweise zwischen den blauen Linien.
- **Homogenität → boxplot()**, ggf. Löschen der Ausreiser
  - Verschiedene Untergruppen (muss nicht Problem sein)
  - Ansonsten lieber robuste Methoden wählen

Punkt- & Parameterschätzung

Schätzproblem

Allgemeines - «best guess»

- (Vektor von) Parameter heisst Klein-Theta θ (θ = λ; θ = (θ_1, θ_2) = (μ, σ))
- Parameter sind Konstanten und keine Zufallsvariablen, zufällig = Stichprobe
- Wahrscheinlichkeitsaussagen über Parameter nicht möglich (ausser Bayes)
- Funktion, die von iid Stichprobe aber nicht von unbekannten θ abhängt nennt man Statistik. Statistiken, die Parametern schätzen = Schätzer / Estimator

Schätzer für Erwartungswert E(X) und Varianz Var(X)

- Arithmetische Mittel als Schätzer für Erwartungswert $\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
- Stichprobenvarianz als Schätzer für Erwartungswert der Varianz σ^2

Eigenschaften von Schätzern

- **Erwartungstreue** – Keinen systematischen **Bias** (Verzerrung)
- **Konsistenz** – mit wachsendem n muss MSE und Bias verschwinden Muss asymptotisch erwartungstreu
- $MSE_\theta(T) = Var_\theta(T) + \left(Bias_\theta(T)\right)^2$
- **Effizienz** $MSE_\theta(T_1) \leq MSE_\theta(T_2)$
- **Äquivarianz** – muss gewisse Transformationen v. Schätzer mitmachen
- **Robust** – Immun gegen Ausreiser
- **Effiziente Berechenbarkeit**

Bayes-Schätzer (Konstr. Schätzer)

g(θ) = A-Priori | L(θ) = Likelihood-Fun $= \prod_{i=1}^{m} f(x_i|\theta)$ Produkt einz. Dichten

A-Posteriori $= \boldsymbol{h(\theta)} = \frac{\boldsymbol{g(\theta) \cdot L(\theta)}}{\boldsymbol{\int g(\tilde{\theta}) \cdot L(\tilde{\theta})\, d\tilde{\theta}}}$

Parameter θ wird als ZV mit Wahrscheinlichkeitsverteilung interpretiert

Konstruktion von Schätzern

- **Analogie- oder Einsetzprinzip**
- **Momentenmethode**
- **Maximum-Likelihood-Methode**
  - **Analytisch** – Ableitung + Nullstelle
  - **Numerisch**:

1. Gittersuche: expand.grid()
Gitter mit möglichen Werten in Funktion einsetzen und Maximum suchen. Gut bei mehreren lokalen Maxima Mit Log-Likelihood-Funktion **iterieren**

2. Iterativ: optim() / fitdist()
Suche mit Plug-in Schätzer (Gittersuche, Momentenschätzer, Zufallswerte) nach grösster Parameterkombination Effizient, gefährlich bei mehr lok. Max optim() sucht Min! mit -1 multiplizieren fitdist(data = x , distr = "gamma", *fix.arg = list(scale = 1)*) bei Stand-Vert.

Vertrauens-/Konfidenzintervalle

Allgemeines - Bereichsschätzung

- **Ziel:** Ergänzung unserer Punktschätzung mit einer Genauigkeitsangabe
- **α** = **Fehlerwahrscheinlichkeit**, dass gesuchter Wert nicht im Intervall liegt
- **1 - α ≤** bezeichnet **Überdeckungswahrscheinlichkeit** oder **Konfidenzniveau**
- Überdeckungswahrscheinlichkeit ist aufgrund **≤** in der Regel höher als nötig/verlangt
- Wahrscheinlichkeitsaussage bezieht sich auf Intervall, nicht Parameter
- Je schmaler Intervall, umso genauer. Die Grösse wird wie $\frac{1}{\sqrt{n}}$ kleiner. Um ein KI halber Länge zu erhalten, muss Stichprobengrösse deshalb vervierfacht werden
- Liefert ab n = 25 gute Resultate, bei Schiefe oder n < 25 andere Methoden bevorzugen

Modell/Situation	$(1-\alpha)$-Konfidenzintervall
μ einer Normalverteilung σ^2 bekannt	$\left[\bar{X}_n - \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}, \bar{X}_n + \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}\right]$ `library(BSDA)` `z.test(x,sigma.x=σ,conf.level=α)$conf.int`
μ einer Normalverteilung σ^2 unbekannt	$\left[\bar{X}_n - \frac{s_n}{\sqrt{n}} \mathrm{qt}_{(1-\frac{\alpha}{2},(n-1))}, \bar{X}_n + \frac{s_n}{\sqrt{n}} \mathrm{qt}_{(1-\frac{\alpha}{2},(n-1))}\right]$ `t.test(x,conf.level=α)$conf.int`
$E[X]$ für beliebige Verteilung (bei grosser Stichprobe)	$\left[\bar{X}_n - \frac{s_n}{\sqrt{n}} \mathrm{qt}_{(1-\frac{\alpha}{2},(n-1))}, \bar{X}_n + \frac{s_n}{\sqrt{n}} \mathrm{qt}_{(1-\frac{\alpha}{2},(n-1))}\right]$ `t.test(x,conf.level=α)$conf.int`
beliebige asymptotisch normalverteilt Schätzer $\hat{\theta}$	$\left[\hat{\theta} - q_{1-\frac{\alpha}{2}} \sqrt{\widehat{Var(\hat{\theta})}}, \hat{\theta} + q_{1-\frac{\alpha}{2}} \sqrt{\widehat{Var(\hat{\theta})}}\right]$
p einer Binomialverteilung	`binom.test(x,n,conf.level=α)$conf.int`
λ einer Poisson-Verteilung	`poisson.test(x,T,conf.level=α)$conf.int`

KI-Länge l

$n \geq \left(\frac{2\sigma}{l} \cdot q_{1-\frac{\alpha}{2}}\right)^2$

$l = 2\frac{\sigma}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}}$

Bootstrap

- **Keine** Annahme über Verteilungsfunktion.
- Unabhängig von ZGWS
- Zieht aus Daten eine eigene Stichproben **mit** zurücklegen
- Weniger genau als Berechnung mit korrekter Verteilungsannahme
- **boot(data = x, statistic = schaetzer, R = N)**
- **boot.ci(type = "bca")**

ZGWS gilt bei n ≥ 25 und keine Ausreiser / Schiefe / Lang-/ Kurzschwänzigkeiten

Statistische Tests / Hypothesentest

Packages
- boot
- BSDA
- car
- coin
- fitdistrplus

	H_0 beibehalten	H_0 verworfen
H_0 ist wahr	richtige Entscheidung	Fehler 1.Art
H_0 ist falsch	Fehler 2. Art	richtige Entscheidung

p-Wert < α durch puren Zufall

p-Wert nicht klein genug obwohl H_A wahr

iid (independent & identically distributed) gilt für alle Tests!

Allgemeines
- Statistische Tests basieren auf **Falsifikation**
- Die **Ausgangshypothese** H_0 sagt i.d.R. aus, dass der **Effekt**, den wir gerne belegen würden, **nicht existiert**.
- **Hoffen**, dass die Daten **genügend stark gegen** diese **Ausgangshypothese** sprechen und wir sie verwerfen.
- **Hypothesen** betreffen i.d.R. **Parameter** von Verteilungen (p, λ, μ, etc.) | $H_0: \mu = \mu_0 = 200, H_1 = H_A: \mu \neq \mu_0 = 200$ | $H_0: \lambda \leq 3,\ H_A: \lambda > 3$ | $H_0: p \leq 0.3,\ H_A: p > 0.3$
- **Hypothesen** werden aufgrund von **Vorwissen** oder durch Anwendung einer/s **anderen Methode**/Verfahrens **aufgestellt**, **nicht** auf **Basis desselben Datensatzes**!
- **Sammeln Belege gegen** die **Ausgangshypothese**, in dem wir schauen, wie wahrscheinlich oder unwahrscheinlich ein Ergebnis ist, wenn die Ausgangshypothese stimmt. **Wahrscheinlichkeitsaussagen** über **Wahr**- oder **Falschheit** von **Hypothesen** sind **grundsätzlich nicht möglich**. **Fehler 1. Art ist der schlimmere!**

Vorgehen/Kochrezept
- Nullhypothese
- Alternativhypothese
- Signifikanzniveau α
- Teststatistik
- Verwerfungsbereich/p-Wert
- Testentscheid

p-Wert
- Gibt Wahrscheinlichkeit an, bei unterstellter Gültigkeit von H_0 eine Beobachtung zu machen, die genauso stark oder stärker gegen H_0 spricht wie unsere Beobachtung
- Wenn p-Wert kleiner/gleich Signifikanzniveau α, lehnen wir H_0 ab

Diskret | **Einstichproben Test** | **Stetig**

Binomialtest
- Bernoulli-Experiment
- Teststatistik X ~ Bin(n, p)
- **binom.test(x, n, p = H_0)**

Poissontest
- Anzahl Ereignisse pro Zeitspanne/Ort
- Vergleich einer Poisson-Beobachtung mit einer festen bekannten Rate
- Teststatistik X ~ Pois(λ)
- **poisson.test(x, T, r = H_0)**

1-Stichproben Fall
- Wilcoxon Vorzeichen Rangsummen Test
- Para. Test = keine Verteilungsannahme/ZGWS wilcoxsign_test(distribution = "exact") für Bindungen / Test, geht auch für stetige Stichpro.

z-Test / Gausstest
- Sd/Var σ ist **bekannt**
- Normalverteilung/grosse Stichprobe → ZGWS
- $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right); Z = \sqrt{n}\frac{(\bar{X}-\mu_0)}{\sigma_X}$
- z.test(mu = , sigma.x =) Verteilung mit .norm()

t-Test
- Sd/Var σ ist **nicht bekannt**
- Normalverteilung/grosse Stichprobe → ZGWS
- $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$;
- $T = \sqrt{n}\frac{(\bar{X}_n-\mu_0)}{S_n} \sim t_{n-1}$
- **t.test()** Der Funktion ganze Stichprobe übergeben! .t()

Macht (Power)
- Wahrscheinlichkeit Nullhypothese zurückzuweisen, wenn sie wirklich falsch ist. Fehler Typ II zu vermeiden
- Macht ist abhängig von α Alpha, n, Varianz, Abstand zur H_0
- **power.t.test(type = "one.sample", alternative)**
- Vier Werte ergeben Fünften: n, delta, sd, sig.level (α), power

MC Permutations-Test
- Ziehen **ohne** zurücklegen
- **independence_test**(x ~ y, distribution = "exact") para. T.

Zweistichproben Tests | **Mehrfache Tests** | **Multiples Testen**

Gepaarte Stichproben - ZGWS
- Differenz D der gepaarten Werte bilden
- Gleiche Stichprobengrösse n
- Normalverteilung/grosse Stichprobe/ZGWS
- **t.-/z.test() paired = TRUE** bei Übergabe (x, y) | oder (x-y) Übergabe Differenz

Gepaarte Stichproben kein ZGWS
- Wilcoxon Vorzeichen Rangsummen Test
- Para. Test = keine Verteilungsannahme, dafür verminderte Macht! - .signedrank()
- **wilcoxsign_test(distribution = "exact")** für Bindungen → Übergabe mit **x ~ y**
- gepaarte Test sind vorzuziehen!

Ungepaarte Stichproben - ZGWS
- Normalverteilung/grosse Stichprobe/ZGWS
- Identische Sd/Var σ der beiden Stichproben
- Ungleiche Stichprobengrösse n und m
- **t.-/z.test(x, y) → paired = FALSE, Faustregel Faktor 4** → var.equal = TRUE

Ungepaarte Stichproben kein ZGWS
- U-Test / Mann Whitney Rangsummen Test
- **wilcox_test(distribution = "exact")** → Bindungen/para. Test / Übergabe **data.frame**
- verminderte Macht - .wilcox()

Multiples Testen - ZGWS
- Simultanes Testen mehrere Hypothesen
- Fehler 1. Art steigt bei diesem Test
- $m = \binom{Anz\ Gruppen}{Anz\ Variabel} = \frac{k(k-1)}{2}$ = Anz Ver.
- Ausrechnen in R mit choose(G, V)
- $Fehler\ 1.Art = 1 - \left(1 - \frac{\alpha}{100}\right)^{(Anz\ Vergl.)}$
- Stichprobengrösse n k(l)einen Effekt
- **pairwise.t.test(p.adjust.method = "bonf", pool.sd = FALSE)**
- Boferoni-Korrektur vergleicht mit **α/m**, Ausgabewerte Funktion **direkt Vergleich** mit Signifikanzniveau α
- Bonf vermindert Macht, Methode holm +

Chi-Quadrat-Test - χ^2
- Zusammenhang zwischen zwei kategoriellen Variablen → Kontingenztafel mit r Zeilen und s Spalten (**aufpassen!!!**)
- Getestet wird nur auf Bestehen **irgendeines** Unterschieds!
- $T = \chi^2_{(r-1)(s-1)}, df = (r-1)(s-1)$
- **chisq.test(matrix)** → Matrix **ohne** Zeilen-/Spalten-**Total** übergeben! - .chisq()

2-Stichproben Binomialtest (A/B-Test 2 x 2 Matrix)
- Stammen Stichproben aus selben Binomial-Verteilung?
- x: E = Anzahl Erfolge | von | n = Anzahl Versuchen
- **prop.test(x = c(E1, E2), n = c(n1, n2), correct = FALSE, …)**

2-Stichproben Poissontest
- Stammen zwei Beobachtungen selben Poisson-Verteilung mit gleicher Rate?
- **poisson.test(x = c(B1, B2), r = 1, ...)**