Stochastik - Übung 9.4.09

wegener@stochastik.uni-hannover.de

Sprechstunde Do 11 - 12 Uhr ; B406

Aufgabe 1

(a) Es liegt eine Beobachtung $(X_1, \ldots, X_n)$ vor

$$P_\vartheta(X = x) = \frac{1}{\vartheta} \cdot \frac{1}{\vartheta} \cdot \ldots \cdot \frac{1}{\vartheta} \cdot I\left(\max_{1 \le \nu \le n} X_\nu \le \vartheta\right)$$

Das Produkt wird $(1/\vartheta)^n$ wird für möglichst kleines ϑ maximal, andererseits wird Indikatorfunktion für $\vartheta < \max X_\nu$ zu Null. Deshalb ist $\hat{\vartheta}_{ML} := \max_{1 \le \nu \le n} X_\nu$ der ML-Schätzer für ϑ

$P_\vartheta(X_\nu = k) = 1/\vartheta$ für $k \in \{1, \ldots, \vartheta\}$ bedeutet soviel wie $P(X_\nu = k) = \frac{1}{\vartheta} I_{\{1, \ldots, \vartheta\}}(k)$

(b) Für $0 < \varepsilon < 1$ gilt

$$P_\vartheta\left(\left|\max_{1 \le \nu \le n} X_\nu - \vartheta\right| > \varepsilon\right) = P_\vartheta\left(\max_{1 \le \nu \le n} X_\nu \le \vartheta - \varepsilon\right)$$

$$= \left(P(X_1 \le \vartheta - 1)\right)^n = \left(\frac{\vartheta - 1}{\vartheta}\right)^n \xrightarrow{n \to \infty} 0$$

$$\left|\overbrace{\underbrace{\max X_\nu}_{\le \vartheta} - \vartheta}^{\le 0}\right|$$

Für $\varepsilon \ge 1$ geht die Wahrscheinlichkeit nur noch schneller gegen Null

Aufgabe 2: Schreibfehler!

richtig: $P(X=k) = \ldots, \; k \in \mathbb{N}_0, \; \max(r+n-a, 0) \leq k \leq \min(r,a)$

$$P_a(X=0) = \frac{\binom{r}{0}\binom{a-r}{n-0}}{\binom{a}{n}} = \frac{(a-r)!}{\cancel{n!}(a-r-n)!} \cdot \frac{\cancel{n!}(a-n)!}{a!}$$

$$= \frac{(a-r)!}{a!} \frac{(a-n)!}{(a-n-r)!} = \frac{(a-n)(a-n-1)\ldots(a-n-r+1)}{a(a-1)\ldots(a-r+1)} < 1 \quad \forall a \in \mathbb{N}$$

Auf der anderen Seite: Für sehr großes a kann man n vernachlässigen, d.h.

$$\sup_{a \in \mathbb{N}, a \geq n+r} P_a(X=0) = 1,$$ das Supremum wird also nicht angenommen.

es gibt also kein Maximum und damit keinen ML-Schätzer

Aufgabe 3

Zuerst: Likelihoodfunktion aufstellen

$$L(x_1, \ldots, x_n; \vartheta) = \prod_{i=1}^{n} f(x_i; \vartheta)$$

$$\log(L(x_1, \ldots, x_n; \vartheta)) = \sum_{i=1}^{n} \log(f(x_i; \vartheta))$$

$$= \sum_{i=1}^{n} \left[-(x_i - \vartheta) - 2\log(1 + \exp(-(x_i - \vartheta))) \right]$$

$$\underbrace{\frac{d \log L}{d\vartheta}}_{h(\vartheta)} = \sum_{i=1}^{n} \left(1 - 2 \frac{\exp(-(x_i - \vartheta))}{1 + \exp(-(x_i - \vartheta))}\right)$$

$$= n - 2 \sum_{i=1}^{n} \frac{\exp(\vartheta - x_i)}{1 + \exp(\vartheta - x_i)} \overset{!}{=} 0$$

$$\lim_{\vartheta \to -\infty} h(\vartheta) = n$$

$$\lim_{\vartheta \to \infty} h(\vartheta) = -n$$

Über die 2. Ableitung oder „durch hinsehen" ergibt sich $h(\vartheta)$ ist streng monoton fallend. Das heißt, es existiert eine eindeutige Nullstelle. Diese liefert den Wert für den ML-Schätzer.

b) Durch ein numerisches Verfahren erhält man als Nullstelle ungefähr den Wert $-0{,}267$

Stochastik Übung

16. 4. 2009

Kovarianzmethode von Rao:

Es seien

$$\mathcal{E} = \{ d : \mathcal{X} \to \Gamma \text{ messbar} \mid E_\vartheta\, d(\vartheta) = \gamma(\vartheta) \ \forall \vartheta \in \Theta \}$$

$$\mathcal{E}_0 = \{ d_0 : \mathcal{X} \to \mathbb{R} \text{ messbar} \mid E_\vartheta\, d_0(\vartheta) = 0 \ \forall \vartheta \in \Theta \}$$

die Klasse der erwartungstreuen Schätzer für $\gamma(\vartheta)$ bzw. die der sogenannten „Nullschätzer".

Dann gilt

$d \in \mathcal{E}$ mit $E_\vartheta (d^*)^2 < \infty \ \forall \vartheta \in \Theta$ ist UMVU-Schätzer für $\gamma(\vartheta)$

$\Leftrightarrow \forall \vartheta \in \Theta,\ d_0 \in \mathcal{E}_0$ mit $E_\vartheta\, d_0^2(\vartheta) < \infty$ gilt

$$\mathrm{Cov}_\vartheta (d^*(X), d_0(X)) = 0$$
$$\overset{!}{=} E_\vartheta (d^*(X) \cdot d_0(X))$$

Aufgabe 1)

$$X \sim P(\lambda) \Leftrightarrow p(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Schätzer für λ: $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$

$\overline{X}$ ist erwartungstreu: $E\overline{X} = \frac{1}{n} \sum_{i=1}^{n} \underbrace{EX_i}_{=\lambda} = \frac{n\lambda}{n} = \lambda$

Sei nun $d_0(x_1, \dots, x_n)$ ein „Nullschätzer"

Zur Erinnerung: $Ef(X) = \sum_k f(k) \cdot P(X=k)$

$$E_\lambda (d_0(x_1, \dots, x_n) \cdot \overline{X}) = \sum_{x_1=0}^{\infty} \dots \sum_{x_n=0}^{\infty} \left[d_0(x_1, \dots, x_n) \cdot \frac{x_1 + \dots + x_n}{n} \left(\frac{\lambda^{x_1}}{x_1!} e^{-\lambda} \right) \dots \left(\frac{\lambda^{x_n}}{x_n!} e^{-\lambda} \right) \right]$$

$$E_\lambda \overset{!}{=} e^{-\lambda n} \sum_{x_1=0}^{\infty} \cdots \sum_{x_n=0}^{\infty} d_0(X_1 \ldots X_n) \cdot \frac{x_1+\ldots+x_n}{n} \cdot \frac{1}{x_1! \ldots x_n!} \lambda^{x_1+\ldots+x_n}$$

$$\overset{!}{=} e^{-\lambda n} \sum_{k=0}^{\infty} \frac{k}{n} \Bigg[\sum_{\substack{(x_1 \ldots x_n) \in \mathbb{N}^n \\ x_1+\ldots+x_n = k}} d_0(x_1 \ldots x_n) \frac{1}{x_1! \ldots x_n!} \Bigg] \lambda^k = 0$$

wegen $[\,]=0$ für alle λ

Also folgt mit Rao die Behauptung

Wir benutzen nun $E\, d_0(X_1 \ldots X_n) = 0$

$$\underset{\text{wissen wir}}{0 \overset{\uparrow}{=}} E\, d_0(X_1 \ldots X_n) = \underbrace{e^{-\lambda n}}_{>0} \sum_{x_1=0}^{\infty} \cdots \sum_{x_n=0}^{\infty} d_0(x_1 \ldots x_n) \frac{1}{x_1! \ldots x_n!} \underbrace{\lambda^{x_1+\ldots+x_n}}_{>0 \text{ für } \lambda>0}$$

Sortieren nach Potenzen von λ: $= e^{-\lambda n} \sum_{k=0}^{\infty} \underbrace{\Bigg[\sum_{\substack{(x_1 \ldots x_n) \in \mathbb{N}^n \\ x_1+\ldots+x_n = k}} d_0(x_1 \ldots x_n) \frac{1}{x_1! \ldots x_n!} \Bigg]}_{} \lambda^k$

Weil Erwartungswert $=0$ für alle λ muss Ausdruck in $[\,]=0$ sein für alle k

Aufgabe 2

Verteilung von $M = \max_{1 \le \vartheta \le n} X_\vartheta$

$$P_\vartheta(M = k) = P_\vartheta(M \le k) - P_\vartheta(M \le k-1)$$

$$\overset{!}{=} \left(\frac{k}{\vartheta}\right)^n - \left(\frac{k-1}{\vartheta}\right)^n = \frac{1}{\vartheta^n}\left(k^n - (k-1)^n\right),$$

$$k \in \{1, \ldots, \vartheta\}, \ \vartheta \in \mathbb{N}$$

b) $$E_\vartheta(d(M)) = \sum_{k=1}^{\vartheta} \frac{k^{n+1} - (k-1)^{n+1}}{k^n - (k-1)^n} \cdot \frac{1}{\vartheta^n}\left(k^n - (k-1)^n\right)$$

$$\overset{!}{=} \frac{1}{\vartheta^n} \sum_{k=1}^{\vartheta} \left(k^{n+1} - (k-1)^{n+1}\right) = \frac{\vartheta^{n+1}}{\vartheta^n} = \vartheta$$

Terme in der Summe heben sich auf

also $d(M)$ erwartungstreu

Sei $\varphi : \mathbb{N}^n \to \mathbb{R}$ Nullschätzer

$$E_\vartheta\left[\varphi(X_1,\dots,X_n)\cdot\frac{(\max\limits_{1\le\nu\le n} X_\nu)^{n+1}-(\max X_\nu-1)^{n+1}}{(\max X_\nu)^n-(\max(X_\nu-1))^n}\right]$$

$$=\sum_{x_1=1}^{\vartheta}\dots\sum_{x_n=1}^{\vartheta}\varphi(x_1\dots x_n)\frac{(\max x_\nu)^{n+1}-(\max x_\nu-1)^{n+1}}{(\max x_\nu)^n-(\max x_\nu-1)^n}\cdot\frac{1}{\vartheta^n}$$

Benutzen wieder $E\varphi(X_1\dots X_n)=0$

$$0=\sum_{x_1=1}^{\vartheta}\dots\sum_{x_n=1}^{\vartheta}\varphi(x_1\dots x_n)\cdot\frac{1}{\vartheta^n}\quad >0\ \forall\vartheta\in\mathbb{N}$$

$$=\sum_{t=1}^{\vartheta}\left[\sum_{\substack{(x_1\dots x_n\in\mathbb{N})\\ \max x_\nu=t}}\varphi(x_1\dots x_n)\right]$$

Betrachten [] sukzessive für $\vartheta=1,2\dots$, daraus ergibt sich [] $=0$ für alle $t\in\mathbb{N}$

$$=\frac{1}{\vartheta^n}\sum_{t=1}^{\vartheta}\frac{t^{n+1}-(t-1)^{n+1}}{t^n-(t-1)^n}\underbrace{\left[\sum_{\substack{(x_1,\dots,x_n)\in\mathbb{N}\\ \max x_\nu=t}}\varphi(x_1\dots x_n)\right]}_{=0\ \forall t\in\mathbb{N}}=0$$

Behauptung folgt mit Rao

Aufgabe 1

Momentenmethode liefert Schätzer für die Momente einer Verteilung, d.h.

$$EX \overset{!}{=} \frac{1}{n}\sum_{i=1}^{n} X_i \quad , \quad EX^2 \overset{!}{=} \frac{1}{n}\sum_{i=1}^{n} X_i^2 \quad , \ldots$$

Aus der Vorlesung: $\widehat{EX} = \bar{X}$, $\widehat{Var}(X) = s^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X})^2$

Idee: Wir bestimmen die entsprechenden Momente aus der Stichprobe, setzen sie gleich den „wahren" Werten und lösen nach diesen auf.

(a) $\bar{X} \overset{!}{=} E_{(a,b)} X_1 = \int_{-\infty}^{\infty} x \cdot f_{a,b}(x)\, dx = \int_a^b \frac{x}{b-a}\, dx = \frac{1}{b-a} \frac{x^2}{2}\Big|_a^b = \frac{a+b}{2}$

$$s^2 \overset{!}{=} Var_{(a,b)}(X_1) = \frac{(b-a)^2}{12}$$

Formen jetzt so um, dass $a = f_1(\bar{X}, s^2)$, $b = f_2(\bar{X}, s^2)$ ist.

$$\bar{X} = \frac{a+b}{2} \quad , \quad s^2 = \frac{(b-a)^2}{12}$$

$$b = 2\bar{X} - a \qquad \Rightarrow \quad \frac{(2\bar{X} - a - a)^2}{12} = s^2$$

$$\Leftrightarrow \quad 3s^2 = (\bar{X} - a)^2 \quad \Leftrightarrow \quad \bar{X} - a = \pm\sqrt{3s^2}$$

$$\Leftrightarrow \quad \bar{X} = a \pm \sqrt{3s^2} \Rightarrow a = \bar{X} \pm \sqrt{3s^2}$$

Zusammen mit $a < b$ ergibt das

$$a = \bar{X} - \sqrt{3s^2} \quad , \quad b = \bar{X} + \sqrt{3s^2}$$

$\Rightarrow (\bar{X} - \sqrt{3s^2}, \bar{X} + \sqrt{3s^2})$ ist Schätzer für (a,b)

Für die gegebenen Werte erhält man $(0{,}8280, 2{,}9610)$

(b) Wir zeigen, dass $f_{\mu,\sigma}(x)$, d.h. die Dichte der (X_i), symmetrisch um μ ist, dann ist μ der Erwartungswert.

$$f_{\mu,\sigma}(x) = (F_{\mu,\sigma}(x))' = \frac{-1}{\left(1+e^{-\frac{(x-\mu)}{\sigma}}\right)^2} \cdot e^{-\frac{(x-\mu)}{\sigma}} \cdot \frac{-1}{\sigma}$$

$$= \frac{e^{-\frac{(x-\mu)}{\sigma}} \cdot \frac{1}{\sigma}}{\left(1+e^{-\frac{(x-\mu)}{\sigma}}\right)^2} = e^{\frac{2x}{\sigma}}\left(1+e^{-\frac{x}{\sigma}}\right)^2$$

Vergleichen $f_{\mu,\sigma}(\mu+x)$ und $f_{\mu,\sigma}(\mu-x)$

$$f_{\mu,\sigma}(\mu+x) = \frac{e^{-\frac{x}{\sigma}} \cdot \frac{1}{\sigma}}{\left(1+e^{-\frac{x}{\sigma}}\right)^2} \qquad f_{\mu,\sigma}(\mu-x) = \frac{e^{\frac{x}{\sigma}} \cdot \frac{1}{\sigma}}{\left(1+e^{\frac{x}{\sigma}}\right)^2} = \frac{e^{-\frac{x}{\sigma}} \cdot \frac{1}{\sigma}}{\left(1+e^{\frac{x}{\sigma}}\right)^2}$$

Also $f_{\mu,\sigma}(x+\mu) = f_{\mu,\sigma}(x-\mu)$ und damit ist μ der Erwartungswert.

Wissen $\frac{\pi^2}{3} = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f_{\mu,\sigma}(x)\,dx$

$$\operatorname{var}(X_1) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f_{\mu,\sigma}(x)\,dx$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 \cdot \frac{e^{-\frac{(x-\mu)}{\sigma}}}{\left(1+e^{-\left(\frac{x-\mu}{\sigma}\right)}\right)^2}\,dx$$

$$= \int_{-\infty}^{\infty} \sigma^2 y^2 \frac{e^{-y}}{(1+e^{-y})} \cdot \frac{1}{\sigma}\,\sigma\,dy$$

$$= \sigma^2 \cdot E_{0,1}\, X_1^2$$

$$= \sigma^2 \left(\underbrace{\operatorname{Var}_{0,1}(X_1}_{\frac{\pi^2}{3}} - (\overset{0}{E_{0,1} X_1})^2\right) = \frac{\sigma^2 \pi^2}{3}$$

$y = \frac{x-\mu}{\sigma}$

$x = \sigma y + \mu$

$\frac{dx}{dy} = \sigma$

Jetzt ist $\overline{X}$ Schätzer für μ, außerdem bekommen wir durch $S^2 \stackrel{!}{=} \frac{\sigma \pi^2}{3} \Leftrightarrow \frac{3S^2}{\pi^2}$ einen Schätzer für σ^2 bzw $\sqrt{\frac{3S^2}{\pi^2}}$ einen Schätzer für σ

Einsetzen liefert für die gegebenen Werte

$(0{,}6275,\ 1{,}5640)$ als Schätzer für (μ, σ)

Aufgabe 2

(a) zu bestimmen ist das kleinste $X \in \mathbb{R}: F_X(X) \geq \eta$ ist

Für $X \geq 0$

$$\underset{Bin(10, 1/3)}{F(X)} = \sum_{k=0}^{\lfloor x \rfloor} \binom{10}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{10-k}$$

$F(0) = 0{,}01734$

$F(1) = 0{,}10405$

$F(2) = 0{,}29914$

$F(3) = 0{,}5592$

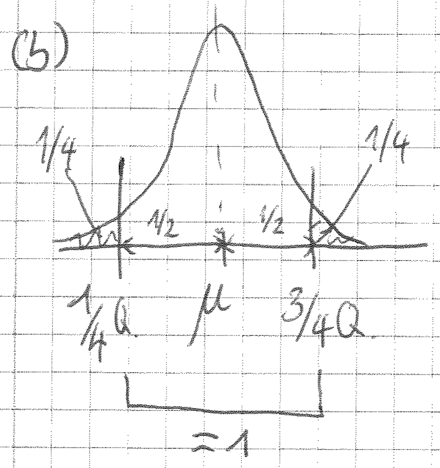
$F(4) = 0{,}78687$

Wegen $F(1) < \frac{1}{4}$, $F(2) \geq \frac{1}{4}$ ist 2 $\frac{1}{4}$ Quartil

" $F(2) < \frac{1}{2}$, $F(3) \geq \frac{1}{2}$ ist 3 $\frac{1}{2}$ Quartil

" $F(3) < \frac{3}{4}$, $F(4) \geq \frac{3}{4}$ ist 4 $\frac{3}{4}$ Quartil

(b)



Wir wollen

$$p\left(\mu - \tfrac{1}{2} \leq X \leq \mu + \tfrac{1}{2}\right) \stackrel{!}{=} \frac{1}{2}$$

$$\Leftrightarrow p\left(\frac{\mu - \frac{1}{2} - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{\mu + \frac{1}{2} - \mu}{\sigma}\right) = \frac{1}{2}$$

$$\phi\left(\frac{1}{2\sigma}\right) - \underbrace{\phi\left(-\frac{1}{2\sigma}\right)}_{1 - \phi\left(\frac{1}{2\sigma}\right)} = 2\phi\left(\frac{1}{2\sigma}\right) - 1$$

$\Leftrightarrow \phi\left((2\sigma)^{-1}\right) = 0{,}75$

$(2\sigma)^{-1} = \xi_{3/4}(0,1)$

$\Leftrightarrow \sigma_0 = \sigma = \frac{1}{2\xi_{3/4}(0,1)}$

Also erfüllt jede $N(\mu, \sigma_0)$-Verteilung für $\mu \in \mathbb{R}$ die geforderte Bedingung

Andere Lösung

$$\xi_{1/4}(\mu, \sigma^2) = \sigma\, \xi_{1/4}(0,1) + \mu$$

$$= -\sigma\, \xi_{3/4}(0,1) + \mu$$

$$\xi_{3/4}(\mu, \sigma^2) = \sigma\, \xi_{3/4}(0,1) + \mu$$

$$\xi_{3/4}(\mu, \sigma^2) - \xi_{1/4}(\mu, \sigma^2) = 2\sigma\, \xi_{3/4}(0,1) \overset{!}{=} 1$$

Nach σ auflösen

Stochastik B 30.4.2009

① Ordnungsstatistiken

	1,9	2,4	2,5	3,0	3,8	4,2	6,3	6,6	7,8	9,2	18,5
F_N	$\frac{1}{11}$	$\frac{2}{11}$	$\frac{3}{11}=0,27$	$\frac{4}{11}$	$\frac{5}{11}$			$\frac{8}{11}=0,73$	$\frac{9}{11}=0,82$		1

$$\frac{1}{n}\sum_{i=1}^{n} X_i = 6,018$$

emp. Median $X_\sigma = 4,2$

emp. Varianz:
$$\left.\begin{array}{l} \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2 = 20,84 \\ \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2 = 22,93 \end{array}\right\} \text{je nach Def.}$$

$$F_{11}(x) = \frac{1}{11}\sum_{i=1}^{n} I(X_i \leq x), \text{ dann ergibt sich}$$

$$\xi_{11,1/4} = 2,5$$
$$\xi_{11,3/4} = 7,8$$

② Wir benötigen die Verteilungsfunktion $F_X(x)$

$$F_X(x) = \int_{-\infty}^{x} f(y)\,dy = \begin{cases} \int_{-\infty}^{x} \frac{1}{2\sigma} e^{\frac{(y-\mu)}{\sigma}}\,dy & \text{für } x \leq \mu \\ \int_{-\infty}^{\mu} \frac{1}{2\sigma} e^{\frac{(y-\mu)}{\sigma}}\,dy + \int_{\mu}^{x} \frac{1}{2\sigma} e^{\frac{-(y-\mu)}{\sigma}}\,dy & \text{für } x>\mu \end{cases}$$

$$= \begin{cases} \frac{1}{2} e^{\frac{x-\mu}{\sigma}}, & x \leq \mu \\ 1-\frac{1}{2} e^{\frac{-x-\mu}{\sigma}}, & x > \mu \end{cases}$$

Verteilungsfunktion ist ~~stetig~~ streng monoton, daher alle Quantile sind eindeutig. Man erkennt wegen $F(\mu)=\frac{1}{2}$, dass μ der Median sein muss

Nach etwas Ausprobieren wählt man als zweites das 0,75 Quantil. Berechnen über Definition

Ansatz: $P(X \le \xi_{0,75}) \ge 0,75$, $P(X \ge \xi_{0,75}) \ge 1-0,75$

$$\Leftrightarrow \quad F(\xi_{0,75}) = 1 - \frac{1}{2} e^{-\frac{(\xi_{0,75}-\mu)}{\sigma}} \ge 0,75$$

$$\frac{1}{2} e^{-\frac{(\xi_{0,75}-\mu)}{\sigma}} \ge 1-0,75$$

$$\Leftrightarrow \quad \mu + \sigma \log(2) \le \xi_{0,75} \le \mu + \sigma \log(2)$$

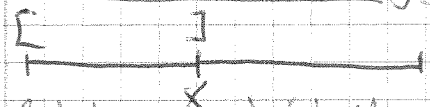
die emp. Quantile $\xi_{0,75} = \mu + \sigma \log(2) \Leftrightarrow \sigma = \frac{\xi_{0,75} - \mu}{\log 2}$

Verwende nun $\xi_{11,\frac{1}{2}}$ und $\xi_{11,\frac{3}{4}}$ als Schätzer für Median und 0,75-Quantil, dies führt für die gegebenen Werte auf (65, 11, 54156) als Schätzer für (μ, σ)

③ a)

$$P(X_{(j)} \le x) = P\Big(\underbrace{\sum_{k=1}^{n} I(X_k \le x)}_{} \ge j\Big)$$

Das beschreibt ein Zufallsexperiment: n-mal werden Werte x „gezogen" mit Wahrscheinlichkeit F(x)



Wertebereich von X. Mit welcher Wahrscheinlichkeit liegt X in $(-\infty, x]$: $P_X((-\infty, x]) = F_X(x)$

liegt ein einzelner gezogener Wert in $(-\infty, x]$. Man ergählt, wie oft dies passiert.

$\Rightarrow \sum_{k=1}^{n} I(X_k \le x) \sim B(n, F(x))$ Binomial-Verteilung

$$= 1 - P\Big(\sum_{k=1}^{n} I(X_k \le x) \le j-1\Big)$$

$$\overset{!}{=} 1 - (n-j+1)\binom{n}{j-1} \int_{F(x)}^{1} t^{j-1}(1-t)^{n-j}\, dt$$

$$= (n-j+1)\binom{n}{j-1}\int_0^{F(x)} t^{j-1}(1-t)^{n-j}\,dt$$

da man die 1 durch $(n-j+1)\binom{n}{j-1}\int_0^1 t^{j-1}(1-t)^{n-j}\,dt$

umschreiben kann

b) Wir zeigen für $\xi :=$ Median von X sind die Anforderung an Median von X_{n+1} erfüllt, wobei $n = 2m+1$

Es ist $F(\xi) \geq \frac{1}{2}$, $1-F(\xi) \geq \frac{1}{2} \Leftrightarrow F(\xi-) \leq \frac{1}{2}$

mit $F(\xi-) = \lim_{\varepsilon \to 0} F(\xi - \varepsilon)$

$$P(X_{(m+1)} \leq \xi) \overset{(a)}{=} (m+1)\binom{2m+1}{m}\int_0^{F(\xi)} t^m(1-t)^m\,dt \geq \cancel{(m+1)}\frac{(2m+1)!}{m!\cancel{(m+1)!}}\int_0^{1/2} t^m(1-t)^m\,dt = \frac{1}{2}$$

$m!$

(letzte Gleichheit)

$$\int_0^{1/2} t^m(1-t)^m\,dt = \frac{1}{2}\,B(m+1, m+1)$$

Beta-Funktion

$$= \frac{1}{2}\,\frac{\Gamma(m+1)\,\Gamma(m+1)}{\Gamma(2m+2)} = \frac{1}{2}\,\frac{m!\,m!}{(2m+1)!}$$

Gamma-Funktion

$$P(X_{(m+1)} < \xi_-) = (m+1)\binom{2m+1}{m}\int_0^{F(\xi-)} t^m(1-t)^m\,dt$$

$$\leq \frac{(2m+1)!}{m!\,m!}\int_0^{1/2} t^m(1-t)^m\,dt = \frac{1}{2}$$

Stochastik – Übung 7.5.09

$$s^2 := \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Aufgabe 1)

Man kann das erwartende Ergebnis mit einer Monte-Carlo-Simulation berechnen / überprüfen.

Wiederhole n Mal:

erzeuge 100 $N(2,3)$ - verteilte ZV

bestimme Median dieser ZV

gilt $|$emp. Median – „wahrer" Median$| < \frac{1}{10}$

dann Zähler erhöhen

ENDE

betrachte $\text{Zähler}/n$, sollte sich für ~~genügend große~~ großes n der der gesuchten Lösung annähern

Manchmal gibt es nur ZV aus einer $N(0,1)$-Verteilung. Wie bekommt man dann $N(2,3)$-verteilte ZV?

$$X \sim N(0,1) \Rightarrow \underset{\sqrt{3}}{\sigma} N + \underset{2}{\mu} \sim N(2,3) \quad (\mu, \sigma^2)$$

$$X \sim N(\mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$$

Aus Vorlesung bekannt:

$$\sqrt{n}(\xi_{n,\mu} - \xi_{\mu}) \xrightarrow{V} N\left(0, \frac{\mu(1-\mu)}{(F'(\xi_\mu))^2}\right)$$

mit $\xi_{n,\mu}$ emp. μ-Quantil

ξ_μ „wahres" μ-Quantil

F Verteilungsfunktion der untersuchten X

hier: $\mu = \frac{1}{2}$, also gilt

$$\sqrt{n}(\xi_{n,1/2} - \xi_{1/2}) \xrightarrow{V} N\left(0, \frac{\frac{1}{2}(1-\frac{1}{2})}{(\varphi_{2,3}(\xi_{\frac{1}{2}}))^2}\right) \text{ mit } \xi_{1/2} = 2$$

Insgesamt also $\sqrt{n}\,(\xi_{n,\frac12} - 2) \xrightarrow{v} N(0, \frac{3}{2}\pi)$

Damit bestimmen wir die gesuchte Wahrscheinlichkeit

$$P\left(|\xi_{100,1/2} - 2| < \frac{1}{10}\right) = P\left(\frac{\sqrt{100}\,|\xi_{100,1/2} - 2|}{\sqrt{\frac{3}{2}\pi}} < \frac{1}{\sqrt{\frac{3}{2}\pi}}\right)$$

$$= P\left(-\sqrt{\frac{2}{3\pi}} < \sqrt{100}\left(\xi_{100,1/2} - 2\right) < \sqrt{\frac{2}{3\pi}}\right)$$

$$= \phi\left(\frac{2}{3\pi}\right) - \phi\left(\frac{-2}{3\pi}\right) = 2\phi\left(\frac{2}{3\pi}\right) - 1 \approx 0{,}355$$

$$\boxed{\phi(-x) = 1 - \phi(x)}$$

Aufgabe 2)

$$I_x(\vartheta) := E_\vartheta\left(\frac{\partial}{\partial\vartheta} \log(f(x|\vartheta))\right)^2$$

(a) $f(x|\vartheta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\vartheta)^2}{2\sigma^2}}$

$$\log(f(x,\vartheta,\sigma^2)) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(x-\vartheta)^2}{2\sigma^2}$$

$$\frac{\partial}{\partial\vartheta}\log(f(x,\vartheta,\sigma^2)) = -\frac{-2(x-\vartheta)}{2\sigma^2} = \frac{x-\vartheta}{\sigma^2}$$

$$E\left(\frac{\partial}{\partial\vartheta}\log(f(x,\vartheta,\sigma^2))\right) = \int_{-\infty}^{\infty}\left(\frac{x-\vartheta}{\sigma^2}\right)^2 \cdot f(x|\vartheta,\sigma^2)\,dx$$

$$= \frac{1}{\sigma^4} \cdot \mathrm{Var}_{\vartheta,\sigma}(X) = \frac{\sigma^2}{\sigma^4} = \sigma^{-2}$$

c) $f(k|\vartheta) = e^{-\vartheta}\frac{\vartheta^k}{k!}$

$$\log(\ldots) = -\vartheta + k\log\vartheta - \log(k!)$$

$$\frac{\partial}{\partial\vartheta}(\ldots) = -1 + \frac{k}{\vartheta} = \frac{k-\vartheta}{\vartheta}$$

$$E_\vartheta(\ldots)^2 = \sum_{k=0}^{\infty}\left(\frac{k-\vartheta}{\vartheta}\right)^2 e^{-\vartheta}\frac{\vartheta^k}{k!} = \frac{1}{\vartheta^2}\cdot\vartheta = \vartheta^{-1}$$

$$I_X(\vartheta) := Var_\vartheta\left(\frac{\partial}{\partial\vartheta}\log f(X,\vartheta)\right) = E\left(\frac{\partial}{\partial\vartheta}\log(f(X,\vartheta))\right)^2 - \underbrace{\left(E\frac{\partial}{\partial\vartheta}\log(f(X,\vartheta))\right)^2}_{=0\ (A1)}$$

Aufgabe 3:

Bei Likelihood-Quotiententests betrachtet man die Testgröße

$$\frac{p_K(x)}{p_H(x)} = \frac{P_{3/4}(X=x)}{P_{1/2}(X=x)} = \frac{\frac{3}{4}\left(1-\frac{3}{4}\right)^x}{\frac{1}{2}\left(1-\frac{1}{2}\right)^x} = 3\cdot 2^{-(x+1)}$$

Der LQ-Test betrachtet Realisierungen dieser Testgröße, Nimmt sie zu große Werte an, wird die Hypothese verworfen.

$$\mathcal{C} = \left\{\left\{X\in\mathbb{N}_0 : 3\cdot 2^{-(x+1)} > c\right\}, c\geq 0\right\}$$
$$= \left\{\left\{X\in\mathbb{N}_0 : -(x+1)\log 2 > \log c - \log 3\right\}, c\geq 0\right\}$$
$$= \left\{\left\{X\in\mathbb{N}_0 : x+1 < \frac{\log 3-\log c}{\log 2}\right\}, c\geq 0\right\}$$

Uns interessieren nur Einschränkungen ~~von~~ für $X(\in\mathbb{N}!)$, nicht konkrete Werte des Bruches

$c\to 0$: Bruch gegen $+\infty$, keine Einschränkungen mehr.

$c\to\infty$: Bruch gegen $-\infty$, kein x erfüllt die Bed. mehr.

Dazwischen alle Werte von c möglich, also alle beliebigen Fälle ~~Wa~~ für X können eintreten

$$= \left\{\left\{X\in\mathbb{N}_0 : X< a\right\}, a\in\mathbb{N}_0\cup\{\infty\}\right\}$$

b) $P_{1/2}(X\leq\infty) = 1$, da $X\in\mathbb{N}_0$. Für $a\in\mathbb{N}$

$$P_{\frac{1}{2}}(X<a) = \sum_{x=0}^{a-1}\frac{1}{2}\left(\frac{1}{2}\right)^x = \frac{1}{2}\,\frac{1-\left(\frac{1}{2}\right)^a}{1-\frac{1}{2}} = 1-\left(\frac{1}{2}\right)^a \geq \frac{1}{2} > \alpha$$

Daher nur für $c=0$ wird die Bedingung $P_{\frac{1}{2}}(X<a)\leq\alpha$ erfüllt.

$P_{1/2}(X \in C)$ ist der Fehler 1. Art, d.h. die Wahrscheinlichkeit, H abzulehnen, obwohl H richtig ist.

Aussage von (b): Nur, wenn der Verwerfungsbereich leer ist, wird H bei keiner Realisierung ablehnen, gibt es einen Test zum Niveau α, $\alpha < \frac{1}{2}$.

Dies ist ein sehr ungünstiges Ergebnis, allerdings sagt das Lemma von Neyman-Pearson, dass es keine „besseren" Tests als LQ-Tests gibt.

Wir müssen für den Fall $X=0$ noch Randomisierung dazunehmen.

also schlagen wir

$$\phi(x) = \begin{cases} 0, & c > 0 \\ \gamma, & c = 0 \end{cases}$$

mit $\gamma = 2\alpha$ als Test vor. Dieser ist „optimal"

Stochastik 14.5.09

Aufgabe 1

Verwenden den einseitigen Binominaltest aus der Vorlesung (siehe Vorlesung)

Testgröße = Anzahl Geheilte

Die Hypothese wird verworfen, falls die Beobachtungen größer als das das $(1-\alpha)$ - Quantil $B(12, \frac{3}{4})_{1-\alpha}$ der $B(12, 3/4)$ - Verteilung, bei Gleichheit erfolgt Randomisierung

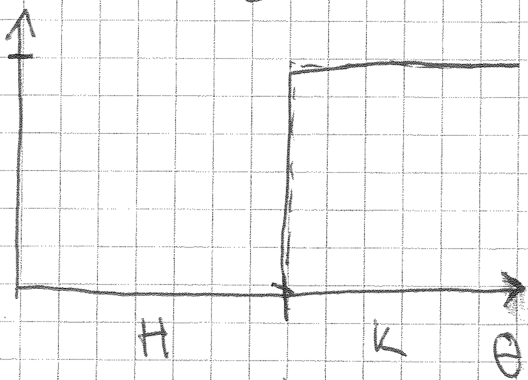
(1) $B(12, 3/4)_{0,9} = 11$, wegen $11 > 7$ ist das Ergebnis mit der Hypothese vereinbar und sie wird nicht verworfen.

(2) $B(12, 3/4)_{0,95} = 11$, wegen $11 = 11$ erfolgt randomisierung mittels einer $B(1, \gamma)$-Verteilten Zufallsvariable, nimmt diese den Wert 1 an, wird H abgelehnt $\gamma = \frac{\alpha - (P(X > 11))}{P(X = 11)} = 0{,}1446$

(3) $B(12, 3/4)_{0,99} = 12$, wegen $12 > 11$ ist die Hypothese mit der Beobachtung vereinbar.

Gütefunktion: $E_\Theta(\varphi(X)) = P_\Theta(X > c) + \gamma P(X = c)$

für $f(x) = \begin{cases} 1 & x > c \\ \gamma & x = c \\ 0 & x < c \end{cases}$



Hier zu beobachten: Ablehngrenze verschiebt sich mit sinkendem α nach rechst bezüglich approx. Gütefkt. sieht man: für $\alpha = 0.01$ versagt die Approximation in der Nähe von 1

Aufgabe 2

Gehen wie in Vorlesung vor und beginnen mit einem Test mit einfacher Hypothese und einfacher Alternative und benutzen dann das Lemma von Neymann-Pearson.

Hier: $f(x,\lambda) = P_\lambda(X_1 = x_1, \dots, X_n = x_n)$

$$= e^{-\lambda n} \frac{1}{x_1! \dots x_n!} \lambda^{x_1 + \dots + x_n}$$

für $x = (x_1, \dots, x_n) \in \mathbb{N}_0^n$ und $\lambda > 0$

Für $\lambda = \lambda_1 > \lambda_0$ und ~~konstante~~ geeignete Konstante k ist

$$\frac{f(x,\lambda_1)}{f(x,\lambda_0)} = e^{-n(\lambda_1 - \lambda_0)} \underbrace{\left(\frac{\lambda_1}{\lambda_0}\right)}_{>1}^{x_1 + \dots + x_n} \left\{\begin{matrix} > \\ = \\ < \end{matrix}\right\} k$$

$$\Longleftrightarrow x_1 + \dots + x_n \left\{\begin{matrix} > \\ = \\ < \end{matrix}\right\} c,$$ da (linke Seite) monoton steigende Fkt. der rechten Seite ist.

Es ist $X_1 + \dots + X_n \sim P(n\lambda)$ für X_i unabhängig gleichmäßig verteilt (i.i.d) $P(\lambda)$, wählen wir $c \in \mathbb{N}_0$ als $(1-\alpha)$-Quantil der $P(n\lambda_0)$-Vert., $\gamma = 0$ im Fall $P_{\lambda_0}(X_1 + \dots + X_n > c) = \alpha$ und $\gamma = \frac{\alpha - P_\alpha(X_1 + \dots + X_n > c)}{P_\alpha(X_1 + \dots + X_n)}$ sonst, so ist

$$\phi(x) = \begin{cases} 1 & \sum x_i > c \\ \gamma & \sum x_i = c \\ 0 & \sum x_i < c \end{cases}$$

ein bester Test zum Niveau $\alpha \in (0,1)$ für $H: \lambda = \lambda_0$, $K: \lambda = \lambda_1$ (Neymann-Pearson)

Da dieser Test nicht vom genauen Wert von λ_1, sondern nur von $\lambda_1 > \lambda_0$ abhängt, ist er auch bester Test für $H: \lambda = \lambda_0$, $K: \lambda > \lambda_0$

Es ist $E_\lambda(\phi(X)) = P_\lambda(X_1 + \dots + X_n > c) + \gamma P_\lambda(X_1 + \dots + X_n = c)$

$$= \sum_{k=c+1}^{\infty} e^{-\mu}\frac{\mu^k}{k!} + \gamma e^{-\mu}\frac{\mu^c}{c!} = g(\mu) \text{ für } \mu := n\lambda$$

Es gilt $g'(\mu) > 0$, also g streng monoton wachsend, d.h. für kleinere Werte von μ sinkt das Niveau nur, d.h. ~~wegen~~ die Grenze α wird auch für $\lambda < \lambda_0$ eingehalten und der Test ist das Gewünschte.

Aufgabe 3

Durch Zentraler Grenzwertsatz (ZGS) ist

$$\tilde{\phi}(X_1, \dots, X_n) = I\left(\frac{X_1 + \dots + X_n - n\lambda_0}{\sqrt{n\lambda_0}} > u_{1-\alpha}\right)$$

ein geeigneter Approximativer Test mit $u_{1-\alpha}$ als $(1-\alpha)$-Quantil der $N(0,1)$-Verteilung

~~Approximation~~:

Gütefunktion: ist $E_\lambda \tilde{\phi}(X_1, \dots, X_n) = \cancel{1 - \Phi(}$

$$= 1 - \Phi\left(u_{1-\alpha} \cdot \sqrt{\frac{n\lambda_0}{n\lambda}} + \frac{n\lambda_0}{\sqrt{n\lambda}} - \frac{n\lambda}{\sqrt{n\lambda}}\right)$$

$$= 1 - \Phi\left(u_{1-\alpha}\sqrt{\frac{\lambda_0}{\lambda}} + \sqrt{n}\,\frac{\lambda - \lambda_0}{\sqrt{\lambda}}\right)$$

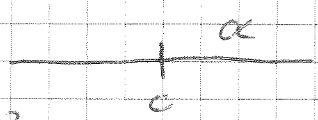
Stochastik B 28.5.09

Blatt 7

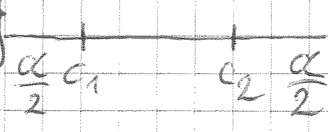
① Fortsetzung von Blatt 6, Aufgabe 2a

Testgröße: $T = \sum_{i=1}^{n} X_i$, „darf nicht zu große Werte annehmen"

Situation hier: Wir testen eine zweiseitige Hypothese, Testgröße „darf weder zu kleine noch zu große Werte" annehmen. Wir teilen das Signifikanzniveau auf „oben und unten" auf.



$c_1 = \max\{x \in \mathbb{N}_0 : P_{\lambda_0}(T < x) \leq \frac{\alpha}{2}\}$



$c_2 = \min\{x \in \mathbb{N}_0 : P_{\lambda_0}(T > x) \leq \frac{\alpha}{2}\}$

$$\gamma_1 := \begin{cases} 0 & , P_{\lambda_0}(T < c_1) = \frac{\alpha}{2} \\ \dfrac{\frac{\alpha}{2} - P_{\lambda_0}(T < c_1)}{P_{\lambda_0}(T = c_1)} & , P_{\lambda_0}(T < c_1) < \frac{\alpha}{2} \end{cases}$$

$$\gamma_2 := \begin{cases} 0 & , P_{\lambda_0}(T > c_2) = \frac{\alpha}{2} \\ \dfrac{\frac{\alpha}{2} - P_{\lambda_0}(T > c_2)}{P_{\lambda_0}(T = c_2)} & , P_{\lambda_0}(T > c_2) < \frac{\alpha}{2} \end{cases}$$

⇒ Testfunktion

$$\varphi(t) = \begin{cases} 1 & t < c_1 \text{ oder } t > c_2 \\ \gamma_1 & t = c_1 \\ \gamma_2 & t = c_2 \\ 0 & c_1 < t < c_2 \end{cases}$$

es gilt $E_{\lambda_0}\varphi(T) = \alpha$

Erwartungswert

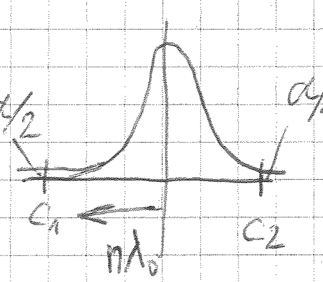
$$E_{\lambda_0}\varphi(T) = 1 \cdot P_{\lambda_0}(T < c_1 \text{ oder } T > c_2) + \gamma_1 P_{\lambda_0}(T = c_1) + \gamma_2 P_{\lambda_0}(T = c_2)$$

$$= P_{\lambda_0}(T < c_1) + P_{\lambda_0}(T > c_2) + \frac{\alpha}{2} - P_{\lambda_0}(T < c_1) + \frac{\alpha}{2} - P_{\lambda_0}(T > c_2)$$

$$= \alpha$$

Für große n Approximation durch Zentraler Grenzwertsatz

$c_1 \approx n\lambda_0 + u_{\frac{\alpha}{2}}\sqrt{n\lambda_0}$



$c_2 \approx n\lambda_0 + u_{\frac{\alpha}{2}}\sqrt{n\lambda_0}$

$\gamma_1 \approx 0$, $\gamma_2 \approx 0$

mit $u_\alpha = \alpha$ − Quantil von $N(0,1)$

Wegen $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$ führt das auf

$$\varphi(T) = \begin{cases} 1, & \frac{|T - n\lambda_0|}{\sqrt{n\lambda_0}} > u_{1-\frac{\alpha}{2}} \\ 0, & \text{sonst} \end{cases}$$

Aufgabe 2

analog zu Blatt 6, Aufgabe 2(a) und Blatt , Aufgabe 1 bzw Vorlesung

---

Blatt 8, Aufgabe 1

Test bei Normalverteilung, $\mu = 0$ vs. $\mu \neq 0$

→ zweiseitiger t-Test

Testgröße $\frac{\sqrt{n}\,|\bar{x}|}{\sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2}}$ hat hier den Wert 0,8368. Das 0,975 Quantil der $\chi^2_{11,\,0{,}975} = 21{,}9201$

also Hypothese mit Beobachtung vereinbar

Aufgabe 2,3,4 entsprechend

Aufgabe 4

Wir führen eine einfache Varianzanalyse durch:

$$X = (X_{11}, \dots, X_{14}, X_{21}, \dots, X_{23}, X_{31}, \dots, X_{33})$$

$$\overline{X}_{i\cdot} := \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} \qquad \overline{X}_{\cdot\cdot} := \frac{1}{n_1+n_2+n_3} \sum_{i=1}^{3} \sum_{j=1}^{n_i} X_{ij}$$

$$H = a_1 = a_2 = a_3$$

$$T(X) = \frac{\frac{1}{2} \sum_{i=1}^{3} n_i \left(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot}\right)^2}{\frac{1}{7} \sum_{i=1}^{3} \sum_{j=1}^{n_i} \left(X_{ij} - \overline{X}_{i\cdot}\right)^2} \approx 2.297 < 4.74 = F_{2,7;0.95}$$

H nicht verwerfen