Betrachten Sie die folgende – bereits geordnete – Stichprobe der Größe $n = 10$:

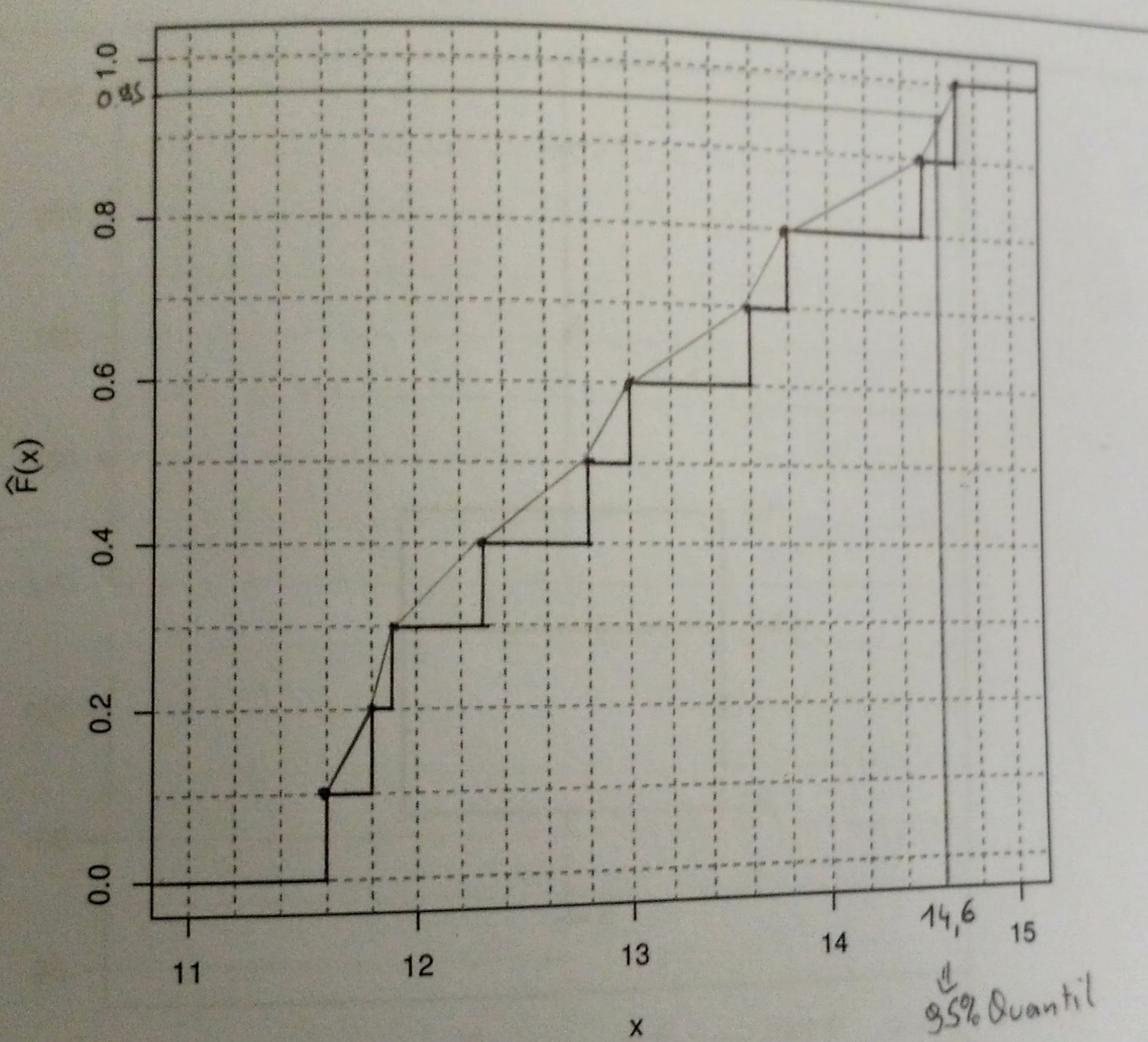
11.6 11.8 11.9 12.3 12.8 13.0 13.6 13.8 14.5 14.7

[2] Zeichnen Sie in die unten stehende Grafik die empirische Verteilungsfunktion.

[1] Bestimmen Sie grafisch das 95%–Quantil vom Typ 4.

[1] Bestimmen Sie den Stichprobenmittelwert.

[1] Bestimmen Sie die Stichprobenvarianz und -streuung.



$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{10} \cdot (11{,}6 + 11{,}8 + \ldots + 14{,}5 + 14{,}7) = 13$$

$$s_n^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 = 1{,}25\dot{3}$$

✓

$$s_n = \sqrt{s_n^2} = 1{,}1195$$

Aufgabe 2

Betrachten Sie die folgende – bereits geordnete – Stichprobe der Größe $n = 19$:

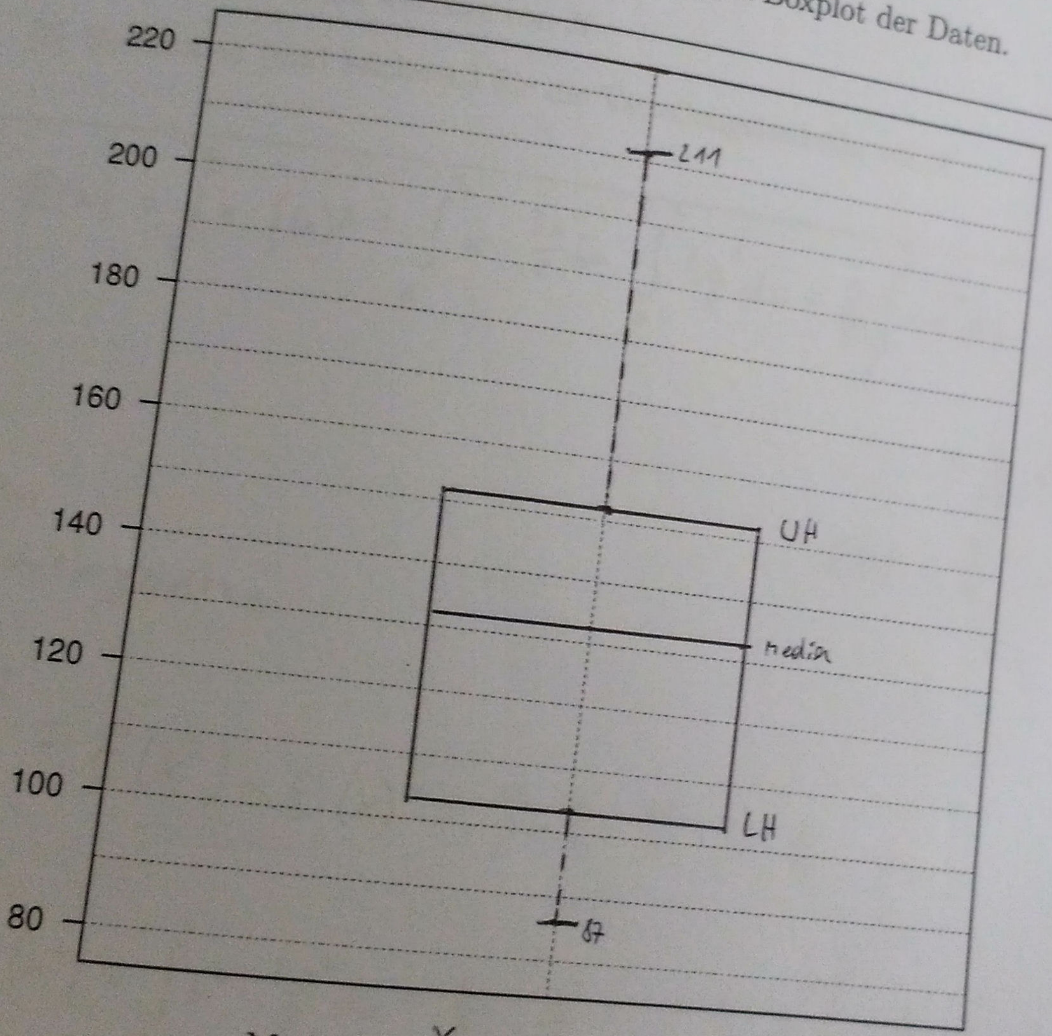
i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$x_{(i)}$	87	87	93	99	103	105	119	129	130	132	138	145	145	152	153	160	180	195	211

[1] Bestimmen Sie den Median.

[1] Bestimmen Sie die Hinges.

[1] Bestimmen Sie auf Basis der Hinges die Fences.

[2] Zeichen Sie in die unten stehende Grafik den Boxplot der Daten.



Median = x_{k+1} falls $n=2k+1$ also ungerade $= x_{10} = 132$

lower Hinge = $\frac{103+105}{2} = 104$

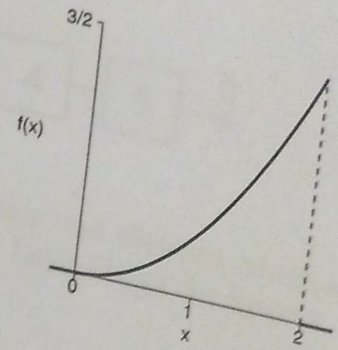
upper Hinge = $\frac{152+153}{2} = 152{,}5$

lower Fence = $Q_1 - 1{,}5\,(Q_3 - Q_1) = 104 - 1{,}5 \cdot (152{,}5 - 104) = 31{,}25$

upper Fence = $152{,}5 + 1{,}5\,(Q_3 - Q_1) = 225{,}25$

Die Dichte einer sG X lautet wie folgt:

$$f(x) = \begin{cases} \dfrac{3x^2}{8} & 0 < x < 2 \\[2ex] 0 & \text{sonst} \end{cases}$$



[1] Bestimmen Sie den Erwartungswert von X.

[2] Bestimmen Sie die Varianz von X.

[2] Bestimmen (und zeichnen) Sie die Verteilungsfunktion von X.

---

a) $\mathbb{E}(X) = \int_0^2 x \cdot f(x)\,dx = \int_0^2 x \cdot \frac{3x^2}{8}\,dx = \int_0^2 \frac{3x^3}{8}\,dx = \frac{3}{8}\int_0^2 x^3\,dx = \frac{3}{8} \cdot \frac{x^4}{4}\Big|_0^2 =$

$= \frac{3}{8} \cdot \frac{16}{4} - \frac{3}{8} \cdot \frac{0}{4}$

$\mathbb{E}(X) = \frac{3}{2}$ ✓

b) Varianz

=> Verschiebungssatz

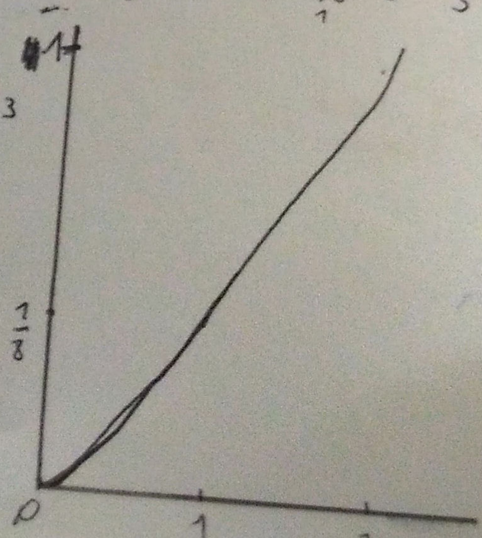
$\sigma_x^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{12}{5} - \frac{9}{4} = \frac{48}{20} - \frac{45}{20} = \frac{3}{20}$ ✓

$\mathbb{E}(X^2) = \int_0^2 x^2 \cdot \frac{3x^2}{8}\,dx = \frac{3}{8} \cdot \int_0^2 x^4\,dx = \frac{3}{8} \cdot \frac{x^5}{5}\Big|_0^2 = \frac{3}{8} \cdot \frac{2^5}{5} = \frac{3}{8} \cdot \frac{32}{5} = \frac{12}{5}$ ✓

c)

$F(x) = \int_0^x f(t)\,dt = \int_0^x \frac{3t^2}{8}\,dt = \frac{3}{8} \cdot \frac{t^3}{3}\Big|_0^x = \frac{x^3}{8}$

$F(x) = \begin{cases} \frac{x^3}{8} & 0 < x < 2 \\ 0 & \text{sonst} \end{cases}$

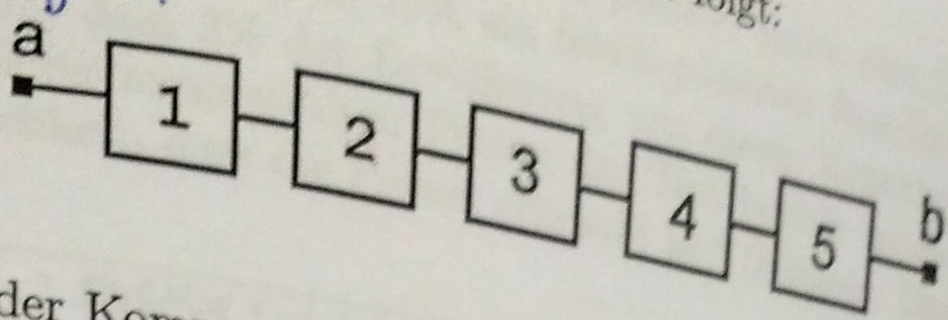


16-3-2016

Die logische Struktur eines Systems sei gegeben wie folgt:

Aufgabe 4

Seriensystem



Die Lebensdauern der Komponenten seien unabhängig und identisch verteilt mit Dichte $f(x) = \frac{1}{10}\, e^{-x/10} I_{(0,\infty)}(x)$ ($\widehat{=}\, \mathsf{Exp}(\tau = 10)$). X sei die Lebensdauer des Systems.

[2] Verteilungsfunktion von X? (Mit Herleitung!) $X = \min\{X_1, X_2, X_3, X_4, X_5\}$

[1] Dichte von X? (Verteilung?)

[2] Erwartungswert, Varianz, Streuung von X?

$$F_{min}(x) = P(X \leq x) = 1 - P(X > x)$$
$$= 1 - P(X_1 > x, X_2 > x, X_3 > x, X_4 > x, X_5 > x)$$
$$= 1 - P(X_1 > x) \cdot P(X_2 > x) \cdot P(X_3 > x) \cdot P(X_4 > x) \cdot P(X_5 > x)$$
$$= 1 - \prod_{i=1}^{5} P(X_i > x)$$
$$= 1 - \prod_{i=1}^{5} (1 - P(X \leq x))$$
$$= 1 - \prod_{i=1}^{5} (1 - (1 - e^{-\frac{1}{10} \cdot x}))$$

$$F_{min}(x) = 1 - e^{-5/10\, x} = 1 - e^{-\frac{1}{2} x}$$ ✓

$$\text{Dichte} = f(x) = F'_{min}(x) = \frac{1}{2} e^{-1/2\, x}$$ (Exponentialverteilung)

E(X)=2, Var(X)=4, Streuung=2

2 [2] Angenommen, ein Test reagiert zu 95% positiv, sollte eine bestimmte Krankheit vorliegen, zeigt aber auch zu 10% ein falsch-positives Resultat. Wenn man davon ausgehen kann, dass 3% der Bevölkerung von dieser Krankheit betroffen ist, und bei einer zufällig ausgewählten Person der Test positiv reagiert, wie groß ist die Wahrscheinlichkeit, dass die Person erkrankt ist? (Hinweis: Bayes'sche Formel)

Ereignisse: K+/K- ... Person hat Krankheit od. nicht
T+/T- ... Test ist positiv/negativ

$$\frac{P(T+|K+)\cdot P(K+)}{P(T+|K+)\cdot P(K+)+P(T+|K-)\cdot P(K-)} = \frac{0{,}95\cdot 0{,}03}{0{,}95\cdot 0{,}03+0{,}1\cdot 0{,}97}$$

$$= 0{,}227 = 22{,}7\%$$

✓

– [1] Ergänzung zu Aufgabe 3: Bestimmen Sie allgemein einen Ausdruck für das p–Quantil x_p $(0 < p < 1)$ der Verteilung. Welchen Wert hat speziell der Median?

=> Skript S. 117

2 [2] Die sG X sei nach P(λ) (Poisson) verteilt mit $\lambda = 9$. Berechnen Sie (approximativ) die Wahrscheinlichkeit $P(7 \leq X \leq 11)$. (Hinweis: Normalapproximation der P(λ)–Verteilung; rechnen Sie mit Stetigkeitskorrektur.)

$\lambda = 9 \Rightarrow$ Approximation nicht ausreichend gut (f. 247 Skript)

$$P(a \leq X \leq b) \simeq \Phi\left(\frac{b+1/2-\lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{a-1/2-\lambda}{\sqrt{\lambda}}\right)$$

$$P(7 \leq X \leq 11) \simeq \Phi\left(\frac{11+1/2-9}{\sqrt{9}}\right) - \Phi\left(\frac{7-1/2-9}{\sqrt{9}}\right) = \Phi(0{,}8\dot{3}) - \Phi(-0{,}8\dot{3})$$

$$= 0{,}7967 - 0{,}2033 = 0{,}5935$$

✓

Eine Stichprobe der Größe $n = 50$ von $X \sim P(\lambda)$ (Poisson) war wie folgt:

$p(x) = P(X=x) = \frac{\lambda^{x} \cdot e^{-\lambda}}{x!}$

x	0	1	2	3	4	5
Häufigkeit	9	17	15	5	3	1

2 [2] Bestimmen Sie allgemein für n Beobachtungen den ML–Schätzer von λ. (Mit genauer Herleitung!)

– [1] ML–Schätzwert von λ auf Basis der gegebenen Stichprobe?

$$L(\lambda) = \prod_{i=1}^{n} p(x_i;\lambda) = \prod_{i=1}^{n} \left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \right)$$

$$\ln L(\lambda) = \sum_{i=1}^{n} \ln\left(\frac{\lambda^{x_i} \cdot e^{-\lambda}}{x_i!} \right) = \sum_{i=1}^{n} \ln(\lambda^{x_i}) + \ln(e^{-\lambda}) - \ln(x_i!) = \sum_{i=1}^{n} x_i \cdot \ln(\lambda) - \lambda - \ln(x_i!)$$

$$= \left(\sum_{i=1}^{n} x_i \cdot \ln(\lambda) - \ln(x_i!)\right) - n \cdot \lambda$$

$$\frac{d \ln L(\lambda)}{d\lambda} = \left(\sum_{i=1}^{n} \frac{x_i}{\lambda}\right) - n$$

$$\left(\sum_{i=1}^{n} \frac{x_i}{\lambda}\right) - n = 0$$

$$\frac{1}{\lambda} \cdot \sum_{i=1}^{n} x_i = n \quad | \cdot \lambda$$

$$\sum_{i=1}^{n} x_i = n\lambda \quad | : n$$

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

2 [2] Bei einer Befragung von 200 wahlberechtigten Personen sagen 104, dass sie vorhaben, für einen bestimmten Kandidaten zu stimmen. Bestimmen Sie ein (approximatives) 90%–Konfidenzintervall für den Wähleranteil dieses Kandidaten.

Bernoulli

$$\hat{p} = \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} x_i \Rightarrow \frac{104}{200} = 0.52 \quad | \; z_{0.95} \hat{=} 1.6449$$

$$\hat{p} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \Rightarrow 0.52 \pm 1.6449 \cdot 0.03533$$

KI: $(0.462, 0.578)$

Für eine Stichprobe x der Größe $m = 8$ von $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$ ergab sich:

```
> data.frame(m=length(x), mean=mean(x), var=var(x), sd=sd(x))
  m  mean   var    sd
  8 11.79 7.687 2.773
```

2

[2] Testen Sie zum Niveau 5%: $\mathcal{H}_0 : \mu_X = 10$ gegen $\mathcal{H}_1 : \mu_X > 10$

$T_0 = \frac{\bar{x}_n - \mu_0}{s_n / \sqrt{n}} = \frac{11.79 - 10}{2.773 / \sqrt{8}} = 1.826$ $\quad t_{7;0.95} = 1.895$

H_0 verwerfen falls: $T_0 \geq t_{n-1;1-\alpha}$

$1.826 < 1.895 \Rightarrow H_0$ nicht verwerfen ✓

Für eine von der obigen Stichprobe unabhängige zweite Stichprobe y der Größe $n = 10$ von $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$ ergab sich:

```
> data.frame(n=length(y), mean=mean(y), var=var(y), sd=sd(y))
   n  mean   var    sd
  10 15.14 10.19 3.193
```

Wenn man davon ausgehen kann, dass $\sigma_Y^2 = \sigma_X^2$:

$S_p^2 = \frac{(8-1) \cdot 7.687 + (10-1) \cdot 10.19}{8 + 10 - 2} = \frac{53.809 + 91.71}{16} = 9.09$

3

[1] Bestimmen Sie den gepoolten Varianzschätzer s_p^2?

[2] Bestimmen Sie ein 95% Konfidenzintervall für $\mu_Y - \mu_X$.

$S_p = \sqrt{9.09} = 3.015$

$t_{16;0.975} = 2.120$

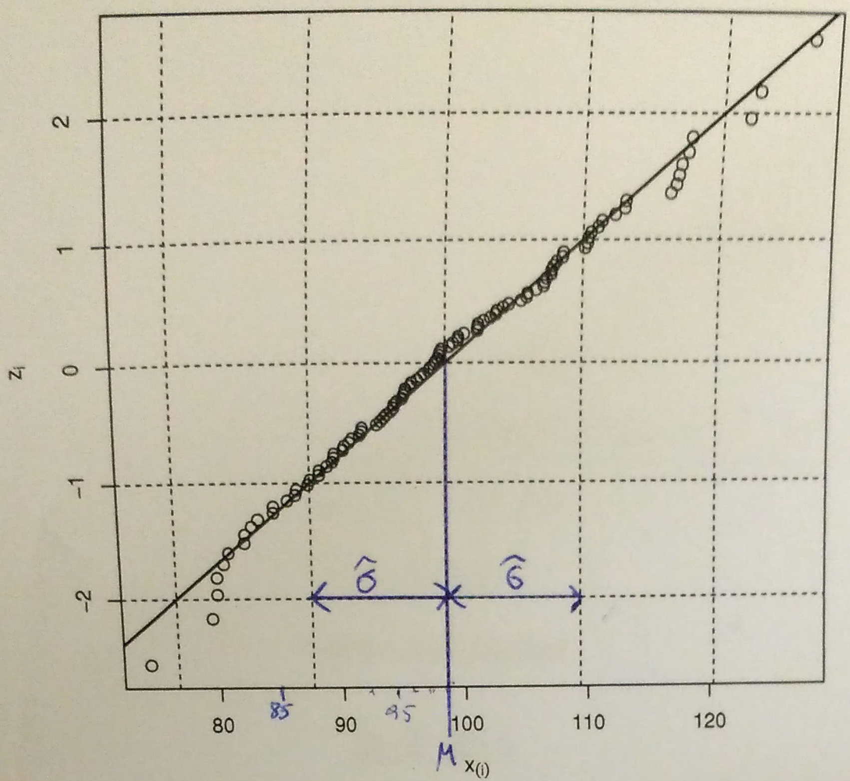
$\bar{X} - \bar{Y} \pm t_{m+n-2;1-\alpha/2} \cdot S_p \sqrt{\frac{1}{m} + \frac{1}{n}}$

$11.79 - 15.14 \pm 2.120 \cdot 3.015 \cdot 0.474$

KI: $(-6.38, -0.32)$ ✓

2

[2] Ein Datensatz der Größe $n = 100$ ergibt im Normal-QQ-Plot das folgende Bild:



Stammen die Daten aus einer Normalverteilung? ☒ ja ☐ nein

Wie groß sind (etwa) Mittelwert $\hat{\mu} \approx 98$ und Streuung $\hat{\sigma} \approx 11$? ✓

3

[3] Stammen die folgenden Beobachtungen aus einer stetigen uniformen Verteilung auf dem Intervall $(0, 1)$? (Testen Sie mit $\alpha = 10\%$.)

Klasse	(0,0.2]	(0.2,0.4]	(0.4,0.6]	(0.6,0.8]	(0.8,1]
Häufigkeit	13	22	14	23	28

Klassen	X	$\hat{p}$	$n\hat{p}$	$(X-n\hat{p})^2/n\hat{p}$
(0,0.2)	13	1/5	20	2.45
(0.2,0.4]	22	1/5	20	0.2
(0.4,0.6]	14	1/5	20	1.8
(0.6,0.8]	23	1/5	20	0.45
(0.8,1]	28	1/5	20	3.2
Summe	100	1	100	8.1

$Q_4 = 8.1$

$\chi^2_{4;0.9} = 7.779$

✓

H_0 verwerfen falls: $Q_{k-1} > \chi^2_{k-1;1-\alpha}$ $\Rightarrow 8.1 > 7.779$ $\Rightarrow$ H_0 wird verworfen