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

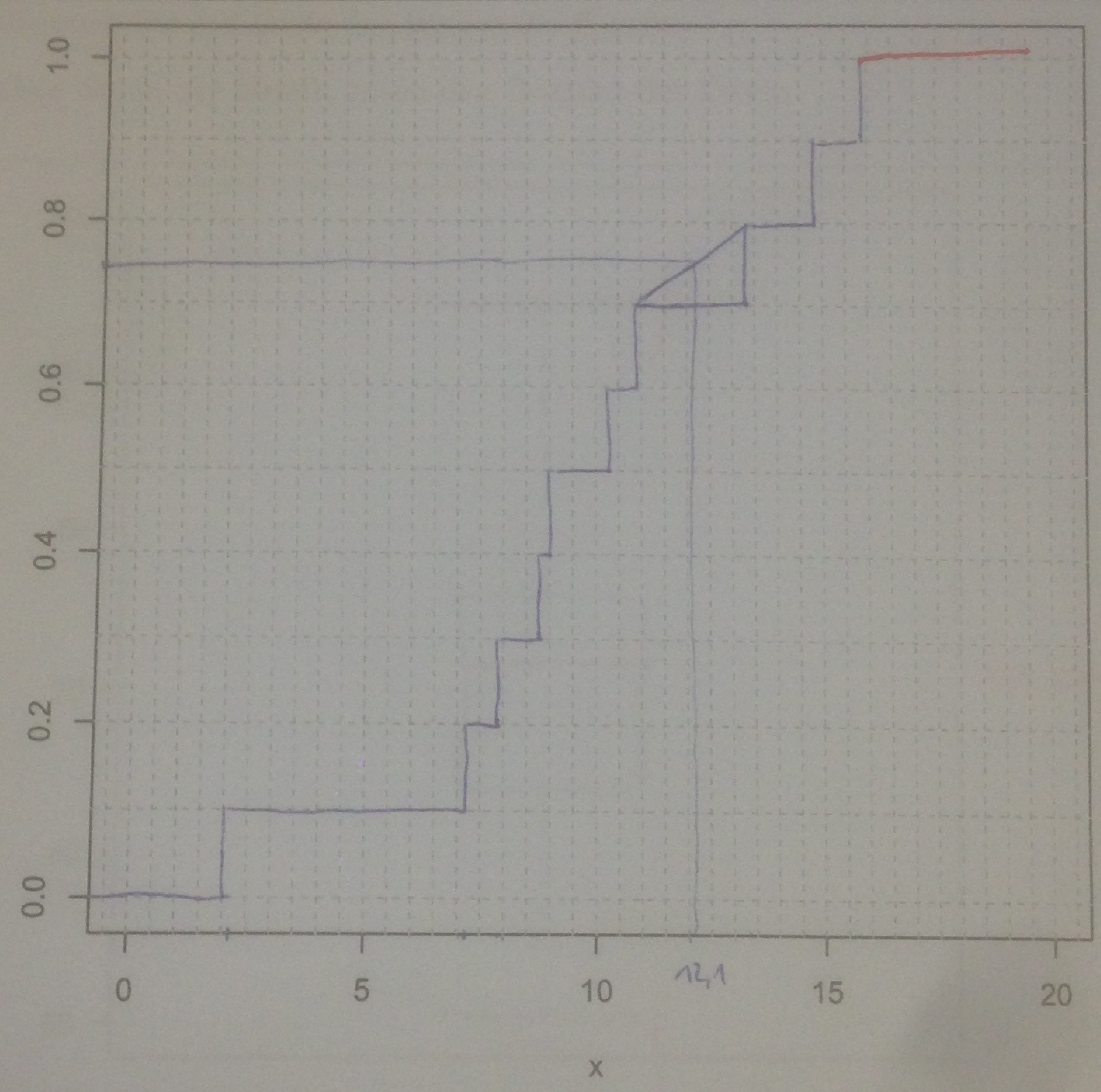
2.2 7.2 7.8 8.7 9.0 10.3 10.9 13.3 14.8 15.8

[2] Zeichnen Sie die empirische Verteilungsfunktion.

[1] Bestimmen Sie grafisch das 75%–Quantil vom Typ 4.

[1] Bestimmen Sie den Stichprobenmittelwert.

[1] Bestimmen Sie die Stichprobenvarianz und -streuung.



75%-Quartil: 12,1

$$\bar{x} = \frac{2.2 + 7.2 + 7.8 + \cdots 15.8}{10} = 10$$

$$s_n^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{1}{9} \cdot 144{,}68 = 16{,}07\dot{5}$$

$$s_n = \sqrt{s_n^2} = 4$$

Betrachten Sie den folgenden (der Größe nach geordneten) Datensatz der Größe $n = 40$:

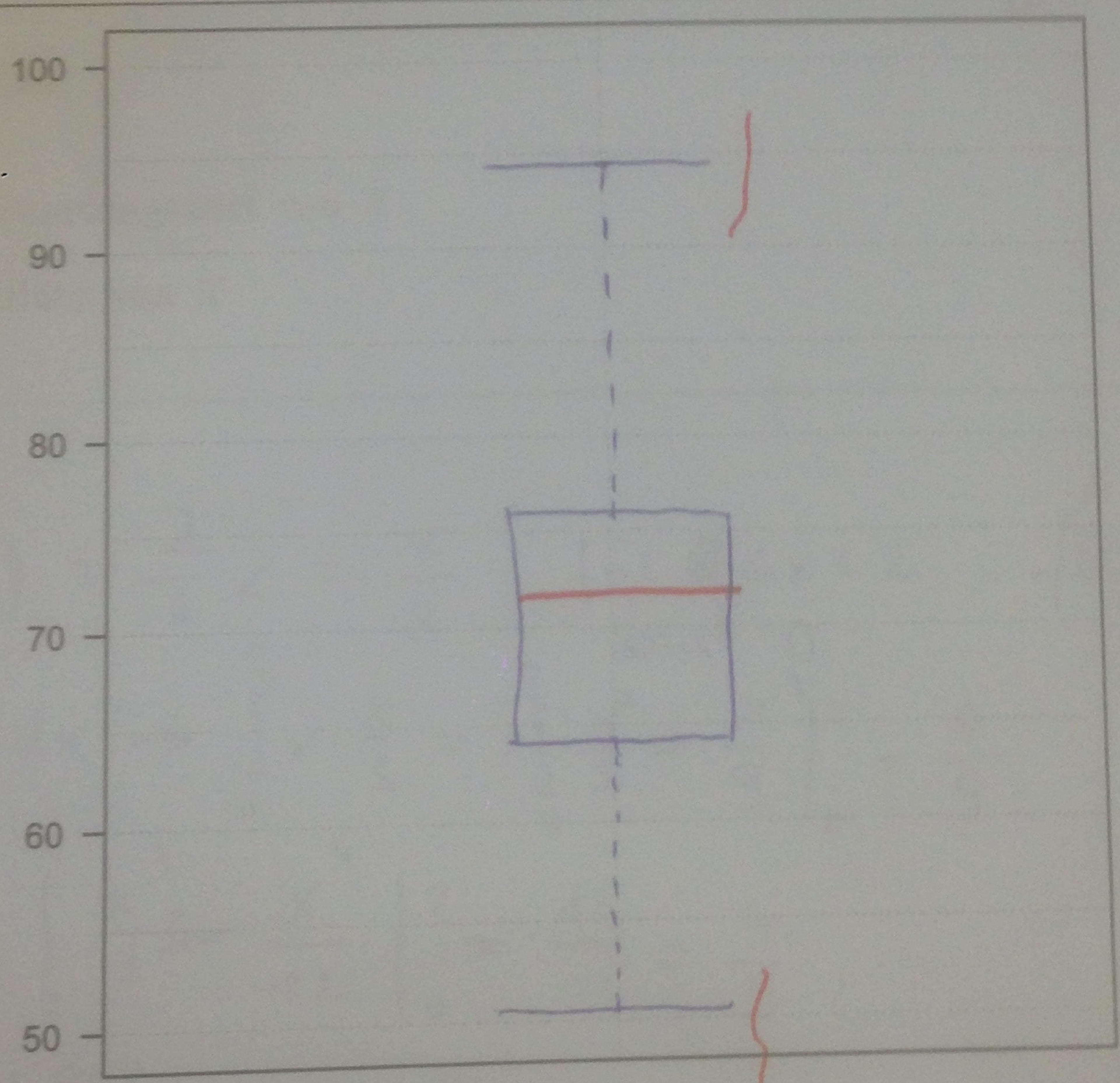
52	56	60	60	60	60	60	60	64	64
64	64	64	64	64	68	68	68	68	72
72	72	72	72	72	72	72	74	76	76
76	76	80	80	84	84	88	88	88	92

[1] Bestimmen Sie den Median.

[1] Bestimmen Sie die Hinges.

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

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



med !

$LF = Q_1 - 1.5(Q_3 - Q_1) =$

$= 64 - 1.5(76 - 64) = 46$

Median = 72

lower Hinge = 64

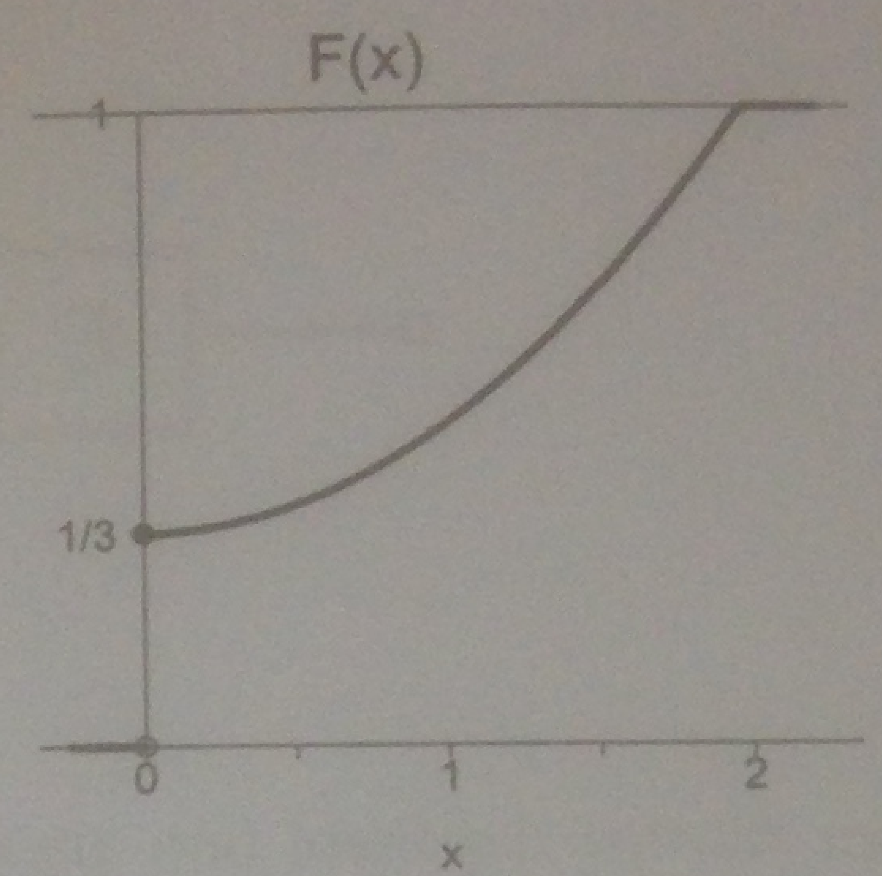
upper Hinge = 76

lower Fence = 46 ✓

upper Fence = 94

Die Verteilungsfunktion einer sG X ist gegeben wie folgt:

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \dfrac{x^2+2}{6} & 0 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$



[1] Handelt es sich um eine ☐ diskrete ☐ stetige ☒ gemischte Verteilung?

Bestimmen Sie:

[2] den Erwartungswert von X

[2] die Varianz von X

$f_x(x) = F'(x) = \frac{2}{6}x = \frac{x}{3}$ für $0 \leq x < 2$
sonst 0

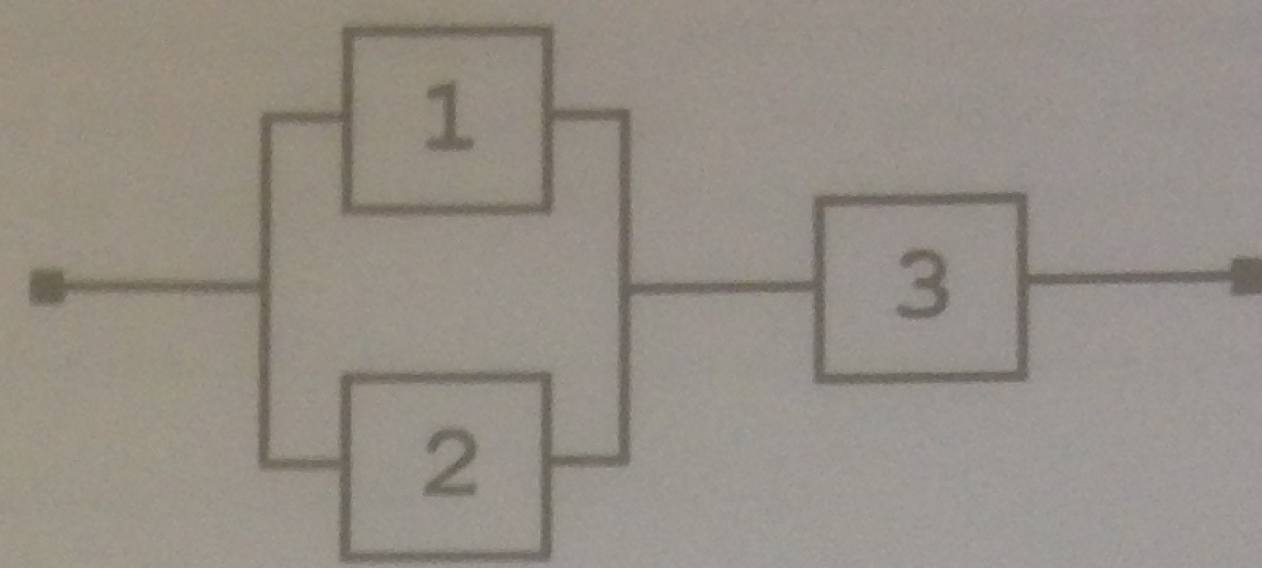
$f(x) = \begin{cases} 0 & x<0 \\ \frac{x}{3} & 0 \leq x < 2 \\ 0 & x>0 \end{cases}$

$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx \rightarrow \int_0^2 x \cdot \frac{x}{3} = \int_0^2 \frac{x^2}{3} = \frac{x^3}{9}\Big|_0^2 = \frac{8}{9}$

$E(X^2) = \int_0^2 x^2 \frac{x}{3}dx = \int_0^2 \frac{x^3}{3}dx = \frac{x^4}{12}\Big|_0^2 = \frac{16}{12} = \frac{4}{3}$

$Var(X) = \frac{4}{3} - \left(\frac{8}{9}\right)^2 = 0{,}5432$

Die logische Struktur eines Systems sei gegeben wie folgt:



Die Lebensdauern der Komponenten seien unabhängig und identisch verteilt mit Dichte $f(x) = e^{-x} I_{(0,\infty)}$ ($\widehat{=}$ Exp(1)). Wenn X die Lebensdauer des Systems ist, bestimmen Sie:

[3] die Verteilungsfunktion von X

[1] die Dichte von X

[1] den Erwartungswert von X

$X_n = \max\{X_1, X_2\}$ $\quad = 1-(2e^{-2y} - e^{-3y}) =$

$= 1 - 2e^{-2y} + e^{-3y}$

Dichte → Ableiten

$\to f(x) = 2e^{-2y} + 3e^{-3y}$

$E(X) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$

$Y_n = \min\{X_1 X_2, X_3\}$

$F_{max}(x) = \prod F_X(x)$

$= (1-e^{-\lambda y})(1-e^{-\lambda y})$

$\Rightarrow (1-e^{-y})^2 = 1-2e^{-y}+e^{-2y}$

$F_{min}(Y) = 1-e^{-n\lambda y}$

$\Rightarrow F_{min}(Y) = 1 - \prod_{i=1}^{n} [1 - F_i(y)]$

$\to 1-((1-e^{-y}))(1-2e^{-y}+e^{-2y}))$

$\to 1-((1-(1-e^{-y}))(1-(1-2e^{-y}+e^{-2y})))$

$= 1-((1-1+e^{-y})(1-\{1+2e^{-y}-e^{-2y}))$

$= 1-(e^{-y}\cdot(2e^{-y}-e^{-2y}))$ ✓

2

[2] Eine Ärztin ist zu 50% davon überzeugt, dass bei einer Person eine bestimmte Erkrankung vorliegt. Zur Abklärung führt sie einen Schnelltest durch, der zwar bei Vorliegen der Krankheit mit Sicherheit positiv reagiert, aber auch zu 10% ein falsch positives Ergebnis liefert. Wenn nun der Test positiv reagiert, wie groß ist danach die Überzeugung der Ärztin? (Hinweis: Bayes'sche Formel)

~~P(T).. Test positiv~~

$P(E^c)$ = P(E)... Person erkrankt = 0,5

$P(T|E) = 1$

$P(T|E^c) = 0{,}1$

$$P(E|T) = \frac{P(T|E) \cdot P(E)}{P(T|E) \cdot P(E) + P(T|E^c) \cdot P(E^c)}$$

✓

$$= \frac{1 \cdot 0{,}5}{1 \cdot 0{,}5 + 0{,}1 \cdot 0{,}5} = 0{,}909$$

→ 90,9 %

–

[1] Wenn der Radius eines Kreises zwischen 1 und 3 stetig uniform verteilt ist, welche Kreisfläche kann man erwarten? (Hinweis: LoTUS)

Fläche = $r^2 \pi$ r zw. 1 und 3 uniform verteilt $\sim U(1,3)$ ✓

$Y = X^2 \pi \rightarrow \cancel{X} = \sqrt{\frac{Y}{\pi}} = g(x)$

$E(Y)$

2

[2] Welche der folgenden R–Commands generiert $n = 100$ unabhängige Realisationen einer exponentialverteilten sG mit Erwartungswert 3? (Hinweis: Inversionsmethode)

τ soll 3 sein

☐
```
u <- runif(100)
x <- -log(1-u)/3
```

☒
```
u <- runif(100)
x <- -3*log(1-u)
```

☐
```
u <- runif(100)
x <- -3*log(3*u)
```

$F_x(x) = 1 - e^{-\lambda x}$

$1 - e^{-\lambda x} = u$

$-u + 1 = +e^{-\lambda x}$

$1 - u = e^{-\lambda x}$

$-\lambda x = \ln(1-u)$

$\lambda x = -\ln(1-u)$

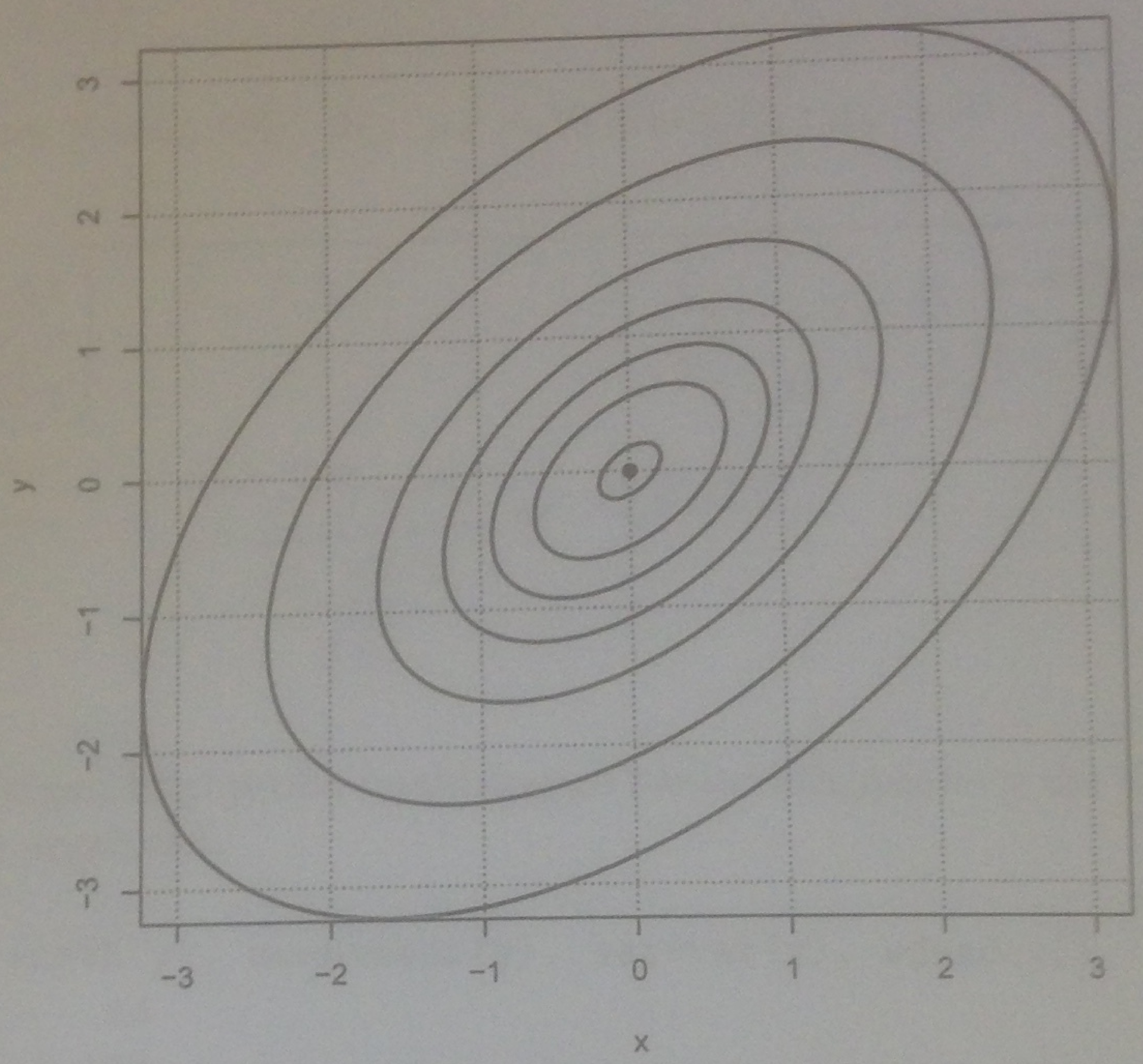
$\rightarrow \frac{1}{\tau} x = -\ln(1-u)$

$\frac{1}{3} x = -\ln(1-u)$ ✓

$x = -3 \ln(1-u)$

[2] Die folgende Abbildung zeigt einige Contourlinien der $N_2(0, 0, 1, 1, 0.5)$ –Verteilung. Zeichnen Sie die beiden Regressionsgeraden ein, d. h., zeichnen Sie $\mathbb{E}(Y|X = x)$ und $\mathbb{E}(X|Y = y)$. (Hinweis: Skriptum/Buch S. 221f)

(S. 221)



[1] X und Y seien unabhängig nach $N(100, 5)$ bzw. $N(120, 10)$ verteilte sGn. Wie ist $D = Y - X$ verteilt?

---

[2] Angenommen, die Bedienungszeit an der Kassa eines Supermarkts ist eine sG mit einem Mittelwert von 3 min und einer Streuung von 2 min. Mit welcher (approximativen) Wahrscheinlichkeit wird für die Bedienung von 16 Kunden insgesamt mehr als 60 min benötigt? (Hinweis: ZGVS)

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

1 [2] Testen Sie zum Niveau 5%: $\mathcal{H}_0 : \mu_X = 10$ gegen $\mathcal{H}_1 : \mu_X > 10$

$\bar{x}_n \pm t$

$t_{n-1;1-\alpha/2} = \cancel{2,365}$

$t_{7;0,975} = 2,365$ 1-seitig!

$$T_0 = \frac{\bar{x}_n - \mu_X}{s_n/\sqrt{n}} = \frac{11,79-10}{\frac{2,773}{\sqrt{8}}} \approx 1,8258$$

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

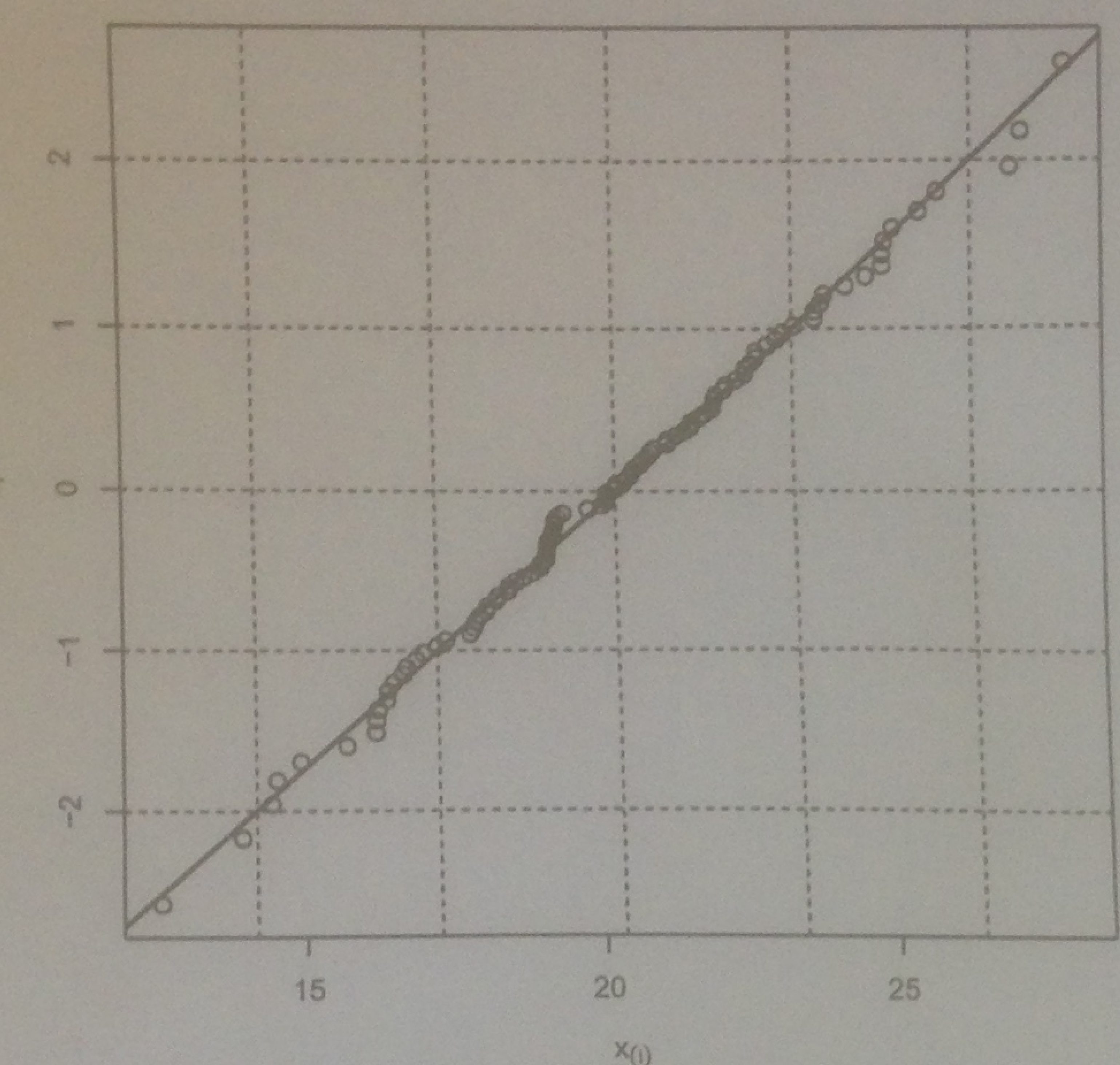
Bestimmen Sie unter der Annahme $\underline{\sigma_Y^2 = \sigma_X^2}$:

[1] den gepoolten Varianzschätzer S_p^2

[2] ein 95% –Konfidenzintervall für $\Delta = \mu_Y - \mu_X$

1

[2] Ein Datensatz der Größe $n = 100$ stellt sich im Normal-QQ-Plot wie folgt dar:



Stammen die Daten aus einer Normalverteilung? ☒ ja ☐ nein

Wie groß sind (etwa) Mittelwert: $\widehat{\mu} \approx 20$ Streuung: $\widehat{\sigma} \approx 5$?

[1] Bei einer Befragung von 300 Personen sprachen sich 45% für einen bestimmten Kandidaten aus. Bestimmen Sie ein 95%–Konfidenzintervall für den Wähleranteil dieses Kandidaten.

---

[2] Bei der Kreuzung von bestimmten Planzen ergeben sich laut Theorie drei Genotypen im Verhältnis 7 : 8 : 5. Ist die Theorie haltbar, wenn bei einem Experiment mit 100 Planzen 42 vom Typ1, 37 vom Typ2 und 21 vom Typ3 sind? ($\alpha = 10\%$)
