STK1110 - Statistiske metoder og dataanalyse høsten 2019
Løsningsforslag til oppgaver i læreboka for uke 39

Vinnie Ko & Ørnulf Borgan

September 25, 2019

Lærebok: *Modern Mathematical Statistics with Applications (2nd ed.)* - Devore & Berk

Exercise 9.4

Husk at feil av type I er definert som "forkast H_0, når H_0 er sann", mens feil av type II er definert som "ikke forkast H_0, når H_a er sann".

Det er "vanlig sedvane" i statistikk at feil av type I svarer til en "alvorlig feil", så man ønsker liten sannsynlighet for feil av type I.

Hvis vi da forkaster H_0 kan vi være "rimelig sikre" på at alternativet H_a er riktig. (Men hvis vi ikke forkaster H_0, kan vi ikke være "rimelig sikre" på at H_0 er riktig. Det kan godt være at H_a er riktig, men at undersøkelsen har ikke nok "beviskraft" til å vise det pga. forholdsvis stor sannsynlighet for feil av type II.)

Når vi skal avgjøre hvilken formulering av alternativet som er relevant, kan vi derfor fokusere på hvilken konklusjon vi vil trekke hvis vi forkaster H_0.

i) Test $H_0 : \mu = 5$ mot $H_a : \mu > 5$, hvor $\mu =$ forventet radioaktivitetsnivå i pCi/L.
Hvis vi her forkaster H_0, kan vi være "rimelig sikre" på at $H_a : \mu > 5$ er sann, dvs. at vannet er utrygt.

ii) Test $H_0 : \mu = 5$ mot $H_a : \mu < 5$, hvor $\mu =$ forventet radioaktivitetsnivå i pCi/L.
Hvis vi her forkaster H_0, kan vi være "rimelig sikre" på at $H_a : \mu < 5$ er sann, dvs. at vannet er trygt.

Så hvis målet med undersøkelsen er å vise at vannet er trygt, må vi velge den siste formuleringen av hypoteseprøvingsproblemet.

Exercise 9.9

a) Forkastningsområdet $R_1 = \{x : x \leq 7 \text{ eller } x \geq 18\}$ er mest egnet, siden det refererer til situasjonen hvor den alternative hypotesen innheholder både verdier at p mindre enn 0.50 og verdier av p større enn 0.50. (Husk at $H_a : p \neq 0.5$ og ikke $H_a : p > 0.5$ eller $H_a : p < 0.5$.)

Vi bruker dette forkastningsområdet i resten av oppgaven.

b) Tabell 1 gir en oversikt over mulige konklusjoner i hypotesetesting. Basert på dette, har vi her at:
- feil av type I: Vi konkluderer at de potensielle abonnentene foretrekker en av to TV-leverandørene ($p \neq 0.5$), mens de potensielle abonnentene ikke har noe preferanse ($p = 0.5$).
- feil av type II: Vi konkluderer at de potensielle abonnentene ikke har noe preferanse for TV-leverandørene ($p = 0.5$), mens de foretrekker en av to TV-leverandørene ($p \neq 0.5$).

	H_0 er sann	H_a er sann
Ikke forkast H_0	Riktig konklusjon $p = 1 - \alpha$	Type II feil $p = \beta$
Forkast H_0	Type I feil $p = \alpha$	Riktig konklusjon $p = 1 - \beta$ ($1 - \beta$ er teststyrken)

Table 1: En oversikt over mulige situasjoner i hypotesetesting.

c) La $X =$"antall personer i utvalget som foretrekker den første leverandøren".
Vi har da at $X \sim \text{Bin}(25, p)$. Når H_0 er sann, er $p = 0.5$.

Fra Tabell 1 har vi at

$$\begin{aligned}\alpha &= P(\text{feil av type I}) = P(X \leq 7 \text{ eller } X \geq 18 \,|\, p = 0.50) \\ &= P(X \leq 7 \,|\, p = 0.50) + P(X \geq 18 \,|\, p = 0.50) \\ &= P(X \leq 7 \,|\, p = 0.50)) + 1 - P(X \leq 17 \,|\, p = 0.50) \\ &= 1 + F_X(7) - F_X(17)\end{aligned}$$

der $F_X(x)$ er den kumulative fordelingen for X.

Ved å bruke `R`, finner vi at sannsynligheten for feil av type I er $\alpha = 0.043$.

```
> 1 + pbinom(7, size=25, prob=0.5) - pbinom(17, size=25, prob=0.5)
[1] 0.04328525
```

d) Fra Tabell 1 har vi for $p \neq 0.50$ at

$$\begin{aligned}\beta(p) &= P(\text{feil av type II} \,|\, p) = P(8 \leq X \leq 17 \,|\, p) \\ &= P(X \leq 17 \,|\, p) - P(X \leq 7 \,|\, p) \\ &= F_X(17) - F_X(7).\end{aligned}$$

Ved å bruke `R`, regner vi ut $\beta(p)$ for $p = 0.3, 0.4, 0.6$ og 0.7.

```
> pbinom(17, size=25, prob=0.3) - pbinom(7, size=25, prob=0.3)
[1] 0.4881334
>
> pbinom(17, size=25, prob=0.4) - pbinom(7, size=25, prob=0.4)
[1] 0.8452428
>
> pbinom(17, size=25, prob=0.6) - pbinom(7, size=25, prob=0.6)
[1] 0.8452428
>
> pbinom(17, size=25, prob=0.7) - pbinom(7, size=25, prob=0.7)
[1] 0.4881334
```

Når $p = 0.3$ eller 0.7, er sannsynligheten for feil av type II 48.8%.
Når $p = 0.4$ eller 0.6, er sannsynligheten for feil av type II 84.5%

e) Siden $6 \in R_1$, forkaster vi H_0 på signifikansnivå $\alpha = 4.3\%$.

Exercise 9.11

a) $H_0 : \mu = 10$ mot $H_a : \mu \neq 10$

b) Vi vil rekalibrere vekten hvis vi forkaster H_0, dvs. hvis $\overline{x} \geq 10.1032$ eller $\overline{x} \leq 9.8968$.

Å rekalibrere vekten når det ikke er nødvendig, svarer til en feil av type I.

Fra Tabell 1 har vi at

$$\begin{aligned}
\alpha &= P(\text{feil av type I}) = P(\text{forkast } H_0 | H_0 \text{ er sann}) \\
&= P(\overline{X} \geq 10.1032 \text{ eller } \overline{X} \leq 9.8968 \,|\, \mu = 10) \\
&= P(\overline{X} \geq 10.1032 \,|\, \mu = 10) + P(\overline{X} \leq 9.8968 \,|\, \mu = 10) \\
&= 1 - P(\overline{X} \leq 10.1032 \,|\, \mu = 10) + P(\overline{X} \leq 9.8968 \,|\, \mu = 10) \\
&= 1 - F_{\overline{X}}(10.1032) + F_{\overline{X}}(9.8968).
\end{aligned}$$

Merk at for $\mu = 10$ har vi at $X_i \sim N\left(10, 0.2^2\right)$. Dermed er $\overline{X} = \frac{1}{25}\sum_{i=1}^{25} X_i \sim N\left(10, \frac{0.2^2}{25}\right)$ under H_0.

Ved å bruke R, finner vi $\alpha = 1 - F_{\overline{X}}(10.1032) + F_{\overline{X}}(9.8968) = 0.01$.

```
> 1 - pnorm(10.1032, mean=10, sd = 0.2/5) + pnorm(9.8968, mean=10, sd = 0.2/5)
[1] 0.009880032
```

c) Å ikke rekalibrere vekten når det er nødvendig, svarer til en feil av type II.

Fra Tabell 1 har vi for $\mu \neq 10$ at

$$\begin{aligned}
\beta(\mu) &= P(\text{feil av type II} \,|\, \mu) = P(\text{ikke forkast } H_0 \,|\, \mu) \\
&= P(9.8968 < \overline{X} < 10.1032 \,|\, \mu) \\
&= P(\overline{X} \leq 10.1032 \,|\, \mu) - P(\overline{X} \leq 9.8968 \,|\, \mu) \\
&= F_{\overline{X}}(10.1032) - F_{\overline{X}}(9.8968).
\end{aligned}$$

Ved å bruke R, finner vi $\beta(\mu)$ for $\mu = 10.1$ og $\mu = 9.8$.

```
> pnorm(q=10.1032, mean=10.1, sd = 0.2/5) - pnorm(q=9.8968, mean=10.1, sd = 0.2/5)
[1] 0.5318812
>
> pnorm(q=10.1032, mean=9.8, sd = 0.2/5) - pnorm(q=9.8968, mean=9.8, sd = 0.2/5)
[1] 0.007760254
```

Når $\mu = 10.1$, har vi $\beta = 0.53$.
Når $\mu = 9.8$, har vi $\beta = 0.0078$.

d) Vi se nå på $z = \dfrac{\overline{x} - 10}{\sigma/\sqrt{n}}$.

Vi har da at

$$\begin{aligned}
& z \geq c \;\text{ eller }\; z \leq -c \\
\iff \quad & \frac{\overline{x} - 10}{\sigma/\sqrt{n}} \geq c \;\text{ eller }\; \frac{\overline{x} - 10}{\sigma/\sqrt{n}} \leq -c \\
\iff \quad & \overline{x} \geq 10 + c\frac{\sigma}{\sqrt{n}} \;\text{ eller }\; \overline{x} \leq 10 - c\frac{\sigma}{\sqrt{n}}.
\end{aligned}$$

Dette svarer til forkastningsområdet i punkt a hvis $c\dfrac{\sigma}{\sqrt{n}} = 0.1032$. Altså får vi at

$$c\frac{0.2}{\sqrt{25}} = 0.1032 \Rightarrow c = 2.58.$$

e) Når $n = 10$ og nullhypotesen er riktig, har vi at $Z = (\overline{X} - 10)/(\sigma/\sqrt{n})$ er standardnormalfordelt.

Da er

$$P(Z \geq z_{\frac{\alpha}{2}} \text{ eller } Z \leq -z_{\frac{\alpha}{2}}) = 1 - \alpha.$$

Så for $\alpha = 0.05$, må vi erstatte c i punkt d med $c = z_{0.025} = 1.96$ (og vi må sette $n = 10$).

f) Basert på verdien av c fra punkt e, er nå forkastningsområdet gitt ved

$$\begin{aligned} &\overline{x} \geq 10 + c\frac{\sigma}{\sqrt{n}} \text{ eller } \overline{x} \leq 10 - c\frac{\sigma}{\sqrt{n}} \\ \iff \quad &\overline{x} \geq 10 + 1.96\frac{0.2}{\sqrt{10}} \text{ eller } \overline{x} \leq 10 - 1.96\frac{0.2}{\sqrt{10}} \\ \iff \quad &\overline{x} \geq 10.1240 \text{ eller } \overline{x} \leq 9.8760 \end{aligned}$$

Ved å bruke `R` finner vi at $\overline{x} = 10.0203$.

```
> weight.data = c(9.981, 10.006, 9.857, 10.107, 9.888, 9.728, 10.439, 10.214, 10.190, 9
    .793)
> mean(weight.data)
[1] 10.0203
```

Siden $\overline{x} = 10.0203$ ikke ligger i forkastningsområdet, forkaster vi ikke H_0.

g) Av resultatet i punkt d har vi at

$$\overline{x} \geq 10.1032 \text{ eller } \overline{x} \leq 9.8968 \qquad \iff \qquad z \geq 2.58 \text{ eller } z \leq -2.58$$

Exercise 9.19

a) Vi har $n = 16$, $\overline{x} = 94.32$ og $\sigma = 1.20$.

Vi skal teste $H_0 : \mu = 95$ mot $H_0 : \mu \neq 95$ med signifikansnivå $\alpha = 0.01$.

Her er σ kjent, så forkaster H_0 hvis (jf. side 438 i læreboka)

$$z = \frac{\overline{x} - 95}{\sigma/\sqrt{n}} \geq z_{\frac{\alpha}{2}} \text{ eller } z = \frac{\overline{x} - 95}{\sigma/\sqrt{n}} \leq -z_{\frac{\alpha}{2}}$$

Med $\alpha = 0.01$ har vi at $z_{\frac{\alpha}{2}} = z_{0.005} = 2.58$.

Vi finner nå at

$$z = \frac{\overline{x} - 95}{\sigma/\sqrt{n}} = \frac{94.32 - 95}{1.20/\sqrt{16}} = -2.27$$

Siden $-2.58 < z < 2.58$ forkaster vi ikke H_0 på 1% signifikansnivå.

b) Fra Tabell 1 har vi at

$$\begin{aligned} \beta(94) &= P(\text{feil av type II} \,|\, \mu = 94) = P(\text{ikke forkast } H_0 \,|\, \mu = 94) \\ &= P\left(-2.58 < \frac{\overline{X} - 95}{1.20/\sqrt{16}} < 2.58 \,|\, \mu = 94\right) \\ &= P\left(95 - 2.58\frac{1.20}{\sqrt{16}} < \overline{X} < 95 + 2.58\frac{1.20}{\sqrt{16}} \,\middle|\, \mu = 94\right) \\ &= P\left(94.226 < \overline{X} < 95.774 \,\middle|\, \mu = 94\right) \\ &= P(\overline{X} \leq 95.774 \,|\, \mu = 94) - P(\overline{X} \leq 94.226 \,|\, \mu = 94) \\ &= F_{\overline{X}}(95.774) - F_{\overline{X}}(94.226). \end{aligned}$$

Merk at når $\mu = 94$, har vi $X_i \sim N\left(94, 1.2^2\right)$. Da er $\overline{X} = \frac{1}{16}\sum_{i=1}^{16} X_i \sim N\left(94, \frac{1.2^2}{16}\right)$.

Ved å bruke R, finner vi at $\beta(94) = F_{\overline{X}}(95.09) - F_{\overline{X}}(93.55) = 0.224$.

```
> pnorm(95+qnorm(0.995)*1.20/4, mean=94, sd = 1.20/4) - pnorm(95-qnorm(0.995)*1.2/4,
    mean=94, sd = 1.2/4)
[1] 0.224374
```

c) Vi er i samme situasjon som i punkt b, men nå ønsker vi å finne n som gir $\beta(94) = 0.1$. Vi kan da bruke formelen på side 441 i læreboka:

$$n = \left(\frac{\sigma\left(z_{\frac{\alpha}{2}} + z_{\beta}\right)}{\mu_0 - \mu_1}\right)^2 = \left(\frac{1.2\left(2.58 + 1.28\right)}{95 - 94}\right)^2 = 21.43$$

Vi trenger 22 observasjoner.

Exercise 9.25

Det er gitt at IQ-verdiene av alle elevene i første klasse i USA er $N(100, 15^2)$-fordelt.

Vi antar at IQ-verdiene av elevene i første klasse på den aktuelle skolen er $N(\mu, 15^2)$-fordelt.

Vi vil teste $H_0 : \mu = 100 \quad \text{mot} \quad H_a : \ \mu \neq 100$.

For IQ-verdiene gitt i oppgaven er gjennomsnittet $\overline{x} = 116$.

```
> IQ.data = c(107, 113, 108, 127, 146, 103, 108,118, 111, 119)
> mean(IQ.data)
[1] 116
```

Her er $\sigma = 15$ kjent, så vi forkaster H_0 hvis (jf. side 438 i læreboka)

$$z = \frac{\overline{x} - 100}{\sigma/\sqrt{n}} \geq z_{\frac{\alpha}{2}} \ \text{ eller } \ z = \frac{\overline{x} - 100}{\sigma/\sqrt{n}} \leq -z_{\frac{\alpha}{2}}$$

Med $\alpha = 0.05$ har vi at $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$.

Vi finner nå at

$$z = \frac{\overline{x} - 100}{\sigma/\sqrt{n}} = \frac{119 - 95}{15/\sqrt{10}} = 3.37$$

Siden $z > 1.96$ forkaster vi H_0 på 5% signifikansnivå.

Vi konkluderer at IQ-verdiene av elevene i første klasse på den aktuelle skolen er signifikant forskjellig fra det nasjonale gjennomsnittet.

Exercise 9.32

a) Vi har målinger fra $n = 12$ radon detektorer. Vi antar at målingene er $N(\mu, \sigma^2)$-fordelte der både μ og σ er ukjente parametere

Vi vil teste $H_0 : \mu = 100 \quad \text{mot} \quad H_a : \ \mu \neq 100$.

Her er σ ukjent, så vi forkaster H_0 hvis (jf. side 444 i læreboka)

$$t = \frac{\overline{x} - 100}{s/\sqrt{n}} \geq t_{\frac{\alpha}{2},n-1} \ \text{ eller } \ t = \frac{\overline{x} - 100}{s/\sqrt{n}} \leq -t_{\frac{\alpha}{2},n-1}$$

Med $n = 12$ og $\alpha = 0.05$ har vi at $t_{\frac{\alpha}{2},n-1} = t_{0.025,11} = 2.20$.

For dataene gitt i oppgaven finner vi $\overline{x} = 98.375$ og $s = 6.109$.

```
> radon.data = c(105.6, 90.9, 91.2, 96.9, 96.5, 91.3, 100.1, 105, 99.6, 107.7, 103.3, 92
    .4)
> mean(radon.data)
[1] 98.375
> sd(radon.data)
[1] 6.109475
```

Dermed er

$$t = \frac{\overline{x} - 100}{s/\sqrt{n}} = \frac{98.375 - 100}{6.109/\sqrt{12}} = -0.921$$

Siden $-2.20 < t < 2.20$ forkaster vi ikke H_0 på 5% signifikansnivå.

b) Vi ønsker å finne n som gir at sannsynligheten for feil av type II er $\beta = 0.1$ for alternativet $\mu = 95$. I disse beregingene bruker vi verdien $\sigma = 7.5$ for standadardavviket.

Vi kan bruke R-kommandoen `power.t.test` for å bestemme n, jf. R-script til forelesningen 16. september:

```
> power.t.test(n=NULL,delta=95-100,sd=7.5,power=0.90,type="one.sample",alternative="
    two.sided")

     One-sample t test power calculation

              n = 25.6399
          delta = 5
             sd = 7.5
      sig.level = 0.05
          power = 0.9
    alternative = two.sided
```

Vi ser at vi må ha $n = 26$.

Alternativt kan vi som en tilnærming bruke formelen på side 441 i læreboka (som bruker normalfordelingen som en tilnærming til t-fordelingen)

$$n = \left(\frac{\sigma \left(z_{\frac{\alpha}{2}} + z_\beta \right)}{\mu_0 - \mu_1} \right)^2 = \left(\frac{7.5 \left(1.96 + 1.28 \right)}{100 - 95} \right)^2 = 23.64$$

Med denne tilnærmingen finner vi $n = 24$

Exercise 9.39

a) Siden det blir tatt et tilfeldig utvalg av $n = 1000$ bøker, har vi at $X =$ "antall bøker som er plassert feil" er bin(n, p)-fordelt, der p er andelen bøker som er plassert feil.

Vi vil teste $H_0 : \ p = 0.02$ mot $H_a : \ p < 0.02$. Hvis H_0 bli forkastet, er en "rimelig sikker" på at $p < 0.02$ slik at en kan utsette å lage ny inventarliste.

Vi forkaster H_0 hvis (jf. side 451 i læreboka)

$$z = \frac{\widehat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} = \frac{\widehat{p} - 0.02}{\sqrt{0.02 \cdot 0.98/1000}} \leq -z_\alpha.$$

Med $\alpha = 0.05$ har vi at $z_\alpha = z_{0.05} = 1.645$.

En fant at $x = 15$ bøker var plassert feil. Det gir $\widehat{p} = x/n = 15/1000 = 0.015$ og

$$z = \frac{0.015 - 0.02}{\sqrt{0.02 \cdot 0.98/1000}} = -1.13.$$

Siden $z > -1.645$ forkaster vi ikke H_0 på 5% signifikansnivå.

Vi vil anbefale at bibliotekaren lager en nye inventarliste.

b) Vi er her interessert i å finne sannsynligheten for feil av type II hvis $p = 0.01$.

Vi bruker formelen på side 452 i læreboka og får:

$$\begin{aligned}\beta(0.01) &= P(\text{feil av type II} \,|\, p = 0.01) \\ &= 1 - \Phi\left(\frac{p_0 - p - z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) \\ &= 1 - \Phi\left(\frac{0.02 - 0.01 - 1.645\sqrt{0.02(1-0.02)/1000}}{\sqrt{0.01(1-0.01)/1000}}\right) \\ &= 1 - \Phi(0.863) = 0.194\end{aligned}$$

c) Her er vi interessert i å finne sannsynligheten for at vi forkaster H_0 hvis $p = 0.05$. Vi er altså interessert i å finne teststyrken $\gamma(p) = P(\text{forkast } H_0 \,|\, p)$ når $p = 0.05$.

Vi har at $\gamma(p) = 1 - \beta(p)$, der $\beta(p)$ er gitt ved formelen på side 452 i læreboka (selv om det ikke er en feil av type II å ikke forkaste når $p = 0.05$).

Altså har vi at

$$\begin{aligned}\gamma(0.05) &= 1 - \beta(0.05) \\ &= \Phi\left(\frac{0.02 - 0.05 - 1.645\sqrt{0.02(1-0.02)/1000}}{\sqrt{0.05(1-0.05)/1000}}\right) \\ &= \Phi(-5.41) \approx 0.\end{aligned}$$

Det er altså tilnærmet null sannsynliget for at en vil utsette å lage en ny invetarliste hvis $p = 0.05$.

Exercise 9.40

$X =$ "antall svarte som blir valgt ut" er bin(n, p)-fordelt, der $n = 1050$ og p er sannynligheten for at et tilfeldig valgt jurymedlem er svart.

Vi vil teste nullhypotesen $H_0 :\ p = 0.25$ mot den alternative hypotesen $H_a :\ p < 0.25$.

Vi forkaster H_0 hvis (jf. side 451 i læreboka)

$$z = \frac{\widehat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{\widehat{p} - 0.25}{\sqrt{0.25 \cdot 0.75/1050}} \leq -z_\alpha.$$

Med $\alpha = 0.01$ har vi at $z_\alpha = z_{0.01} = 2.33$.

Vi har nå at $\widehat{p} = 177/1050 = 0.169$, så vi får

$$z = \frac{0.169 - 0.25}{\sqrt{0.25 \cdot 0.75/1050}} = -6.06$$

Siden $z = -6.06 < -2.33 = -z_{0.01}$ forkaster vi H_0.
Det betyr at vi er rimelig sikre på at det er diskriminering.