2. domača naloga

Statistika 2 in Matematična statistika

Zala Jamšek, 27122040

28. avgust 2014

1 Naloga 1

(a)

Parameter je procent Amričano, ki preferirajo smrtno kazen. Naj bo X_i Bernullijeva spremenljivka

$$X_i \sim \begin{pmatrix} 0 & 1 \\ (1-p) & p \end{pmatrix}$$

Torej je v tem primeru X_i i-ti izmed 537 anketiranih Američanov. V primeru, da na vprašanje odgovori, da je bolj primerna kazen smrt, je vrednost spremenljivke $X_i = 1$, sicer pa je vrednost spremenljivke $X_i = 0$.

Naj bo $X = \sum_{i=1}^{n} X_i$ skupno število Američanov, ki se jim ustrezna kazen zdi smrtna kazen. X je Binomska slučajna spremenljivka in velja:

$X \sim (n, p)$, $E(X) = np$, $D(X) = p(1-p)$. Parameter p, delež Amričanov, ki preferirajo smrtno kazen, je torej:

$$p = \frac{E(X)}{n}.$$

Cenilka $\hat{p} = \frac{X}{n}$ je nepristranska cenilka za p, saj je $E(p) = E(\frac{X}{n}) = \frac{E(X)}{n}$.
Po Centralnem limitnem izreku sledi:

$$\frac{X - E(X)}{\sqrt{D(X)}} \sim N(0,1)$$

$E(X) = np$, $D(X) = np(1-p)$, torej :

$$\begin{aligned} \frac{X - nE(X_i)}{\sqrt{nD(X_i)}} &\sim N(0,1) \\ X &\sim N(np, np(1-p)) \end{aligned}$$

Porazdelitev za $\hat{p} = \frac{X}{n}$ je dobljena tako, da se upanj v zgornjem izrazu deli z n, disperzijo pa z N^2:

$$\begin{aligned} \hat{p} &\sim N(p, \frac{p(1-p))}{n}) \\ \frac{\hat{p} - p}{\sqrt{\frac{p(1-p))}{n}}} &\sim N(0,1) \end{aligned}$$

(b)

Iz zgornjega izraza sledi interval zaupanja, kjer je ϕ, kvantil standardne normalne porazdelitve, α pa stopnja značilnosti:

$$P\left(-\phi(1-\frac{\alpha}{2}) \leq \frac{\hat{p}-p}{\sqrt{\frac{p(1-p))}{n}}} \leq \phi(1-\frac{\alpha}{2})\right) \approx (1-\alpha)$$
$$P\left(\hat{p}-\phi(1-\frac{\alpha}{2})\sqrt{\frac{p(1-p))}{n}} \leq p \leq \hat{p}+\phi(1-\frac{\alpha}{2})\sqrt{\frac{p(1-p))}{n}}\right) \approx (1-\alpha)$$

Ker disperzija ni znana, je p zamenjan s $\hat{p}$:

$$P\left(\hat{p}-\phi(1-\frac{\alpha}{2})\sqrt{\frac{\hat{p}(1-\hat{p}))}{n}} \leq p \leq \hat{p}+\phi(1-\frac{\alpha}{2})\sqrt{\frac{\hat{p}(1-\hat{p}))}{n}}\right) \approx (1-\alpha)$$

```
> n<- 537
> smrt <- 0.52
> jeca <- 0.48
> alpha <- 0.05
> lb <- smrt - qnorm(1-alpha/2)*sqrt(smrt*(1-smrt)/n)
> ub <- smrt + qnorm(1-alpha/2)*sqrt(smrt*(1-smrt)/n)
> lb

[1] 0.4777445

> ub

[1] 0.5622555
```

Interval zaupanja je torej

$$p \in [0.4777445, 0.5622555]$$

(c)

Ker interval zaupanja zajema tudi vrednosti, ki so manjše od $0,5$, ne moremo trditi, da večina (več kot 50%) Američanov podpira smrtno kazen. V primeru, da bi celoten interval zaupanja ležal nad $0,5$ bi to lahko trdili.

(d)

Če ostane odstotek Američanov, ki so za smrtno kazen konstanten, torej $\hat{p} = 0.52$, potem bo velikost intervala zaupanja enaka 2%, pri stopnji značilnosti $\alpha = 2\%$:

$$\begin{aligned} 2\phi(1-0,01)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &\leq 0.02 \\ \frac{\hat{p}(1-\hat{p})}{n} &\leq (\frac{0.01}{\phi(0,99)})^2 \\ n &\geq \frac{\hat{p}(1-\hat{p})}{(\frac{0.01}{\phi(0,99)})^2} \end{aligned}$$

```
> n <- smrt*(1-smrt)/(0.01/qnorm(0.99))^2
> n
```

```
[1] 13508.09
```

Anketirati bi bilo potrebno 13509 ljudi.

2 Naloga 2

(a)

Verjetnosti, da je avto kripa in da je OK enaki,sta enaki. Pričakovana izguba v primeru, da avto kupimo je:

$$E(L|kupi) = \frac{1}{2} \cdot 12 + 0 = 6$$

Pričakovana izguva, če avta ne kupimo, pa je:

$$E(L|\text{ne kupi}) = \frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 4 = 4$$

Optimalna odločitev je, da avta ne kupimo, pričakovana izguba pa je v tem primeru 4.

(b)

Če je servisna knjižica popolna, potem je $y = y_1$. V tem primeru je posteriorna verjetnost:

$$P(\theta_i|y = y_1) = \frac{P(\theta_i)P(y = y_1|\theta_i)}{P(y = y_1|\theta_1) \cdot P(\theta_1) + P(y = y_1|\theta_2) \cdot P(\theta_2)}$$

Torej:

$$\begin{aligned} P(\theta = \theta_1|y = y_1) &= \frac{\frac{1}{2} \cdot 0,2}{\frac{1}{2} \cdot 0,2 + \frac{1}{2} \cdot 0,8} = \frac{0,1}{0,1 + 0,4} = 0,2 \\ P(\theta = \theta_2|y = y_1) &= \frac{\frac{1}{2} \cdot 0,8}{\frac{1}{2} \cdot 0,2 + \frac{1}{2} \cdot 0,8} = \frac{0,4}{0,1 + 0,4} = 0,8 \end{aligned}$$

Pričakovana izguba je torej:

$$E(L|kupi) = 12 \cdot 0,2 + 0 = 2,4$$

$$E(L|\text{ne kupi}) = 4 \cdot 0,2 + 4 \cdot 0,8 = 5$$

Z znano informacijo, da je prometna knjižica OK, se odločimo za nakup avtomobila in v tem primeru je pričakovana izguba $2,4$.

(c)

Stanja knjižice so:
(1): y_1 knjizica OK
(2): y_2 knjižica s pomankljivostmi
(3): y_3 ni knjižice

(1)
V primeru da je knjižica OK, je bilo ze v točki (b) izračunano, da je optimalna strategija, da se kupi avto, pričakovana izguba v tem primeru pa je $2,4$.

(2)
Knjižica s pomankljivostmi:

$$\begin{aligned} P(\theta = \theta_1 | y = y_)) &= \frac{\frac{1}{2} \cdot 0,6}{\frac{1}{2} \cdot 0,6 + \frac{1}{2} \cdot 0,1} = \frac{0,3}{0,3 + 0,05} = \frac{6}{7} \\ P(\theta = \theta_2 | y = y_)) &= \frac{\frac{1}{2} \cdot 0,6}{\frac{1}{2} \cdot 0,6 + \frac{1}{2} \cdot 0,1} = \frac{0,05}{0,3 + 0,05} = \frac{1}{7} \end{aligned}$$

Pričakovana izguba v primeru, da je knjižica s pomankljivostmi:

$$E(L|kupi) = 12 \cdot \frac{6}{7} + 0 = \frac{72}{7}$$

$$E(L|\text{ne kupi}) = 4 \cdot 0,3 + 4 \cdot 0,05 = 4$$

Optimalna odločitev je torej, da se avta ne kupi, pričakovana izguba pa je 4.
(3)
Ni knjižice:

$$\begin{aligned} P(\theta = \theta_1 | y = y_3) &= \frac{\frac{1}{2} \cdot 0,2}{\frac{1}{2} \cdot 0,2 + \frac{1}{2} \cdot 0,1} = \frac{0,1}{0,1 + 0,05} = \frac{2}{3} \\ P(\theta = \theta_2 | y = y_3) &= \frac{\frac{1}{2} \cdot 0,1}{\frac{1}{2} \cdot 0,2 + \frac{1}{2} \cdot 0,1} = \frac{0,1}{0,1 + 0,05} = \frac{1}{3} \end{aligned}$$

Pričakovana izguba v primeru, da je knjižica s pomankljivostmi:

$$E(L|kupi) = 12 \cdot \frac{2}{3} + 0 = 8$$

$$E(L|\text{ne kupi}) = 4 \cdot 0,1 + 4 \cdot 0,05 = 4$$

Optimalna odločitev je torej, da avta ne kupimo, pričakovana izguba pa je 4.

Torej se v obeh preostalih primerih ne odločimo za nakup avtomobila, saj je pričakovana izguba v tem primeru manjša.

Informacija o stanju prometne knjižice vpliva na optimalno odločitev kupca. V primeru da te informacije nima, je optimalna odločitev sprejeta zgolj na podlagi apriorne verjetnosti. Z opazovanjem stanja knjižice pa na pričakovano izgubo vpliva posteriorna verjetnost.

Pričakovana izguba v primeru, da kupec avta ne kupi, je vedno enaka 4, ne glede na stanje knjižice. Optimalna odločitev kupca za nakup avtomobila je torej v primeru, ko je $P(\theta|y) < \frac{1}{3}$. Zato je v primeru, ko je knjižica popolna, optimalna odločitev nakup avtomobila, saj je posteriorna verjetnost $P(\theta = \theta_1|y = y_1) = 0,2$ in je pričakovana izguba nižja, kot v primeru ne nakupa avtomobila. V preostalih dveh primerih pa se kupec za nakup avtomobila ne odloči, enako kot v primeru, ko je znana samo apriorno verjetnost.