Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

LØSNINGSFORSLAG
EKSAMEN I ST1101/ST6200 SANNSYNLIGHETSREGNING
Onsdag 6. juni 2012
Tid: 09:00–13:00

Oppgave 1

In a population of people, where the number of men is equal to the number of women, 5% of men and 0.25% of women are colour blind.

a) What is the probability that a randomly chosen person is colour blind?

Solution. Consider the following events:

A – a randomly chosen person is a man (respectively A^c – a randomly chosen person is a woman),

B – a randomly chosen person is colour blind.

Then $P(B|A) = 0.05$, $P(B|A^c) = 0.0025$, $P(A) = P(A^c) = 0.5$. Due to the total probability theorem,

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c) = 0.05 \cdot 0.5 + 0.0025 \cdot 0.5 = 0.02625.$$

b) A randomly chosen person is colour blind. What is the probability that the person is a man?

Solution. Due to the Bayes theorem,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0.05 \cdot 0.5}{0.02625} = 0.9524.$$

Oppgave 2

The radius of a circle is a random variable with the exponential distribution whose expectation is equal to 1.

a) Let X be the area of the circle. Find the cumulative distribution function of X.

Solution. Denote the radius by R. Then $X = \pi R^2$, and therefore

$$F_X(x) = P(X \le x) = P(\pi R^2 \le x) = P(R \le \sqrt{x/\pi}) =$$

$$= F_R(\sqrt{x/\pi}) = 1 - e^{-\sqrt{x/\pi}}.$$

b) Find the probability that the area of the circle is greater than π but less than 4π.

Solution.

$$P(\pi \le X \le 4\pi) = F_X(4\pi) - F_X(\pi) = \frac{1}{e} - \frac{1}{e^2} = \frac{e-1}{e^2}.$$

c) Show that the median of X is equal to $\pi(\ln 2)^2$.

Solution. The median is the solution of the equation

$$1 - e^{-\sqrt{x/\pi}} = 1/2$$

which is evidently $\pi(\ln 2)^2$.

Oppgave 3

Let $X_1, ..., X_n$ be a random sample (independent and identically distributed random variables) from the distribution with the density (uniform distribution)

$$f(x) = \begin{cases} c & \text{for } 0 \le x \le \theta, \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is the unknown parameter, c is a constant (depending of θ).

a) Find c, EX_i, $VarX_i$.

Solution. $c = 1/\theta$ (evidently),

$$EX_i = \frac{1}{\theta}\int_0^\theta x dx = \frac{\theta}{2},$$

$$EX_i^2 = \frac{1}{\theta}\int_0^\theta x^2 dx = \frac{\theta^2}{3},$$

$$VarX_i = EX_i^2 - (EX_i)^2 = \frac{\theta^2}{12}.$$

b) Find the method of moments estimator $\hat{\theta}_1$ of the parameter θ. Is it unbiased?

Solution. Let M_1 be the first empirical moment i.e.

$$M_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}.$$

Then $\hat{\theta}_1$ is the solution of the equation

$$\frac{\theta}{2} = M_1$$

i.e. $\hat{\theta}_1 = 2\bar{X}$. Since $E\hat{\theta}_1 = 2\bar{X} = \theta$, the estimator is unbiased.

c) Find the maximum likelihood estimator $\hat{\theta}_2$.

Solution. Since

$$f(x;\theta) = \frac{1}{\theta} I_{[0,\theta]}(x),$$

the likelihood function is

$$L(\theta; X) = \frac{1}{\theta^n}\prod_{i=1}^{n} I_{[0,\theta]}(X_i) = \frac{1}{\theta^n} I_{[0,\theta]}(\max_i X_i) = \frac{1}{\theta^n} I_{[\max_i X_i,\infty)}(\theta)$$

with the maximum at $\max_i X_i$. Thus

$$\hat{\theta}_2 = \max\{X_1, ..., X_n\}.$$

d) Show that $\hat{\theta}_2$ is biased. Find a sequence of numbers $\{a_n\}$ such that the estimator

$$\hat{\theta}_3 = a_n \hat{\theta}_2$$

is unbiased.

Solution. Denote the distribution function of X_i by $F(x)$, the distribution function and the density of $\hat{\theta}_2$ by $F_\theta(x)$ and $f_\theta(x)$. Then

$$F(x) = \frac{x}{\theta}$$

for $0 \le x \le \theta$ (0 for $x < 0$ and 1 for $x > \theta$), and therefore

$$F_\theta(x) = [F(x)]^n = \frac{x^n}{\theta^n}$$

for $0 \le x \le \theta$,

$$f_\theta(x) = \frac{n}{\theta^n} x^{n-1} I_{[0,\theta]}(x).$$

$$E\hat{\theta}_2 = \frac{n}{\theta^n} \int_0^\theta x^n dx = \frac{n}{n+1}\theta \neq \theta.$$

The estimator is biased.

$$\hat{\theta}_3 = \frac{n+1}{n}\hat{\theta}_2 = \frac{n+1}{n} \max\{X_1, ..., X_n\}$$

is unbiased.

e) Find the relative efficiency of $\hat{\theta}_3$ with respect to $\hat{\theta}_1$. Which estimator is better (more accurate)?

Solution. Both $\hat{\theta}_1$ and $\hat{\theta}_3$ are unbiased. Find $Var\hat{\theta}_1$ and $Var\hat{\theta}_3$.

$$Var\hat{\theta}_1 = 4Var\bar{X} = \frac{4}{n}VarX_i = \frac{\theta^2}{3n}.$$

To find $Var\hat{\theta}_3$ let us find first $Var\hat{\theta}_2$.

$$E\hat{\theta}_2^2 = \frac{n}{\theta^n} \int_0^\theta x^{n+1} dx = \frac{n}{n+2}\theta^2,$$

$$Var\hat{\theta}_2 = E\hat{\theta}_2^2 - (E\hat{\theta}_2)^2 = \frac{n}{(n+2)(n+1)^2}\theta^2.$$

$$Var\hat{\theta}_3 = \frac{(n+1)^2}{n^2}Var\hat{\theta}_2 = \frac{\theta^2}{n(n+2)}.$$

The relative efficience is

$$\frac{Var\hat{\theta}_1}{Var\hat{\theta}_3} = \frac{n+2}{3}.$$

$\hat{\theta}_3$ is essentially better than $\hat{\theta}_1$.