Norges teknisk–naturvitenskapelige universitet
Institutt for matematiske fag

English

TMA4245 STATISTIKK - Proposed Solution

14. august 2008

Oppgave 1

Define

F: A random toaster unit fails.

We know that the probability that a toaster unit failing is

P(F) = 0.03 = p

a) Since the event "first toaster unit fail" is independent of the event " second toaster unit fail" we can write

$$P(\text{fail } 1 \cap \text{fail } 2) = P(\text{fail } 1)P(\text{fail } 2) = p \times p = 0.0009$$

The probability that 2 out of 7 toaster unit fails can be considered as a result of a binomial experiment, and then we can compute it as

$$P(\text{fail 2 out of 7}) = \binom{7}{2} p^2(1-p)^5 = \frac{7!}{5!2!} 0.03^2 0.97^5 = 0.0163$$

Using the definition of conditional probability and the fact that the events "fail 1", "fail 2" and "not fail 3-7" are independent, we have that

$$
\begin{aligned}
P(\text{fail 1} \cap \text{fail 2}|\text{fail 2 out of 7}) &= \frac{P(\text{fail 1} \cap \text{fail 2} \cap \text{fail 2 out of 7})}{P(\text{fail 2 out of 7})} \\
&= \frac{P(\text{fail 1} \cap \text{fail 2} \cap \text{not fail 3-7})}{P(\text{fail 2 out of 7})} \\
&= \frac{p^2(1-p)^5}{\binom{7}{2} p^2 (1-p)^5} = \frac{1}{\binom{7}{2}} = \frac{5!2!}{7!} = 0.0476
\end{aligned}
$$

b) Using the fact that the event "More than 5 for fail " is equivalent to "no fail in 1-5" we write

$$
P(\text{More than 5 for fail}) = P(\text{no fail in 1-5}) = (1-p)^5 = 0.8587
$$

Using the Law of total probability, we can use that

$$
\begin{aligned}
&P(\text{No two consecutive fails among first 5 units}) = \\
&= \sum_{i=0}^{5} P(\text{No two consecutive fails among first 5 units}|\text{i fail}) \times P(\text{i fail among 5 units})
\end{aligned}
$$

Then,

$$
\begin{aligned}
&P(\text{No two consecutive fails among first 5 units}) = \\
&= P(\text{No two consecutive fails among first 5 units}|\text{0 fail}) \times P(\text{0 fail among 5 units}) \\
&+ P(\text{No two consecutive fails among first 5 units}|\text{1 fail}) \times P(\text{1 fail among 5 units}) \\
&+ P(\text{No two consecutive fails among first 5 units}|\text{2 fail}) \times P(\text{2 fail among 5 units}) \\
&+ P(\text{No two consecutive fails among first 5 units}|\text{3 fail}) \times P(\text{3 fail among 5 units}) \\
&+ P(\text{No two consecutive fails among first 5 units}|\text{4 fail}) \times P(\text{4 fail among 5 units}) \\
&+ P(\text{No two consecutive fails among first 5 units}|\text{5 fail}) \times P(\text{5 fail among 5 units}) \\
&= 1 \times (1-p)^5 + 1 \times \binom{5}{1} p(1-p)^4 + \frac{\binom{5}{2} - 4}{\binom{5}{2}} \times \binom{5}{2} p^2(1-p^3) \\
&+ \frac{1}{\binom{5}{3}} \times \binom{5}{3} p^3(1-p)^2 + 0 \times P(\text{4 fail among 5 units}) + 0 \times P(\text{5 fail among 5 units}) \\
&= 0.9964,
\end{aligned}
$$

where we have used that

$$
\begin{aligned}
P(\text{No two consecutive fails among first 5 units}|\text{0 fail}) &= 1\\
P(\text{No two consecutive fails among first 5 units}|\text{1 fail}) &= 1\\
P(\text{No two consecutive fails among first 5 units}|\text{2 fail}) &= \frac{\binom{5}{2}-4}{\binom{5}{2}}\\
P(\text{No two consecutive fails among first 5 units}|\text{3 fail}) &= \frac{1}{\binom{5}{3}}\\
P(\text{No two consecutive fails among first 5 units}|\text{4 fail}) &= 0\\
P(\text{No two consecutive fails among first 5 units}|\text{5 fail}) &= 0
\end{aligned}
$$

c) We know from the exercise that
$P(F_1 \cup F_2) = 0.03,\ P(F_1) = 0.02,\ P(F_2|F_1) = 0.04.$
Question 1:

$$P(F_2 \cap F_1^c) = P(F_1 \cup F_2) - P(F_1) = 0.03 - 0.02 = 0.01$$

Question 2:

$$
\begin{aligned}
P(F_1|F_2) &= \frac{P(F_1 \cap F_2)}{P(F_2)} = \frac{P(F_2|F_1)P(F_1)}{P(F_2 \cap F_1) + P(F_2 \cap F_1^c)}\\
&= \frac{P(F_2|F_1)P(F_1)}{P(F_2|F_1)P(F_1) + P(F_2 \cap F_1^c)} = \frac{0.4 \times 0.02}{0.4 \times 0.02 + 0.01}\\
&= \frac{0.008}{0.018} = 0.444
\end{aligned}
$$

Oppgave 2

$X \sim N(x; a, \sigma^2)$

a) $a = 2.0,\ \sigma^2 = 0.04 \Rightarrow \sigma = 0.2.$

$$P(X \geq 1.9) = P\left(\frac{X-2}{0.2} \geq -0.5\right) = P(Z \geq -0.5) = 0.6915$$

$$P(1.8 \leq X \leq 2.4) = P\left(\frac{1.8-2}{0.2} \leq \frac{X-2}{0.2} \leq \frac{2.4-2}{0.2}\right) = P(-1.0 \leq Z \leq 2.0)$$
$$= P(Z \leq 2.0) - P(Z \leq -1.0) = 0.8185$$

$$P(X \leq 2.4 | X \geq 1.9) = \frac{P(X \leq 2.4 \cap X \geq 1.9)}{P(X \geq 1.9)} = \frac{P(1.9 \leq X \leq 2.4)}{P(X \geq 1.9)} = \ldots = 0.9670$$

b) $V = \frac{(n-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-1)}$, which means that

$$f(v; n-1) = \frac{1}{2^{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})} v^{\frac{n-1}{2}-1} e^{-v/2}, v \geq 0, \tag{1}$$

since (1) is the density for a χ^2 distribution with $n-1$ degrees of freedom. For a χ^2_{n-1} distribution we have that the mean and variance are given by

$$E[V] = n-1 \quad \text{and} \quad Var[V] = 2(n-1)$$

Using the properties of expectation

$$E[V] = E\left[\frac{(n-1)\hat{\sigma}^2}{\sigma^2}\right] = \frac{(n-1)}{\sigma^2}E[\hat{\sigma}^2] = (n-1) \Rightarrow E[\hat{\sigma}^2] = \sigma^2.$$

Hence, $\hat{\sigma}^2$ is an unbiased estimator for σ^2.

$$Var[V] = Var\left[\frac{(n-1)\hat{\sigma}^2}{\sigma^2}\right] = \frac{(n-1)^2}{\sigma^4}Var[\hat{\sigma}^2] = 2(n-1) \Rightarrow Var[\hat{\sigma}^2] = \frac{2\sigma^4}{n-1}$$

c)

$$P(\chi^2_{n-1,0.95} \leq V \leq \chi^2_{n-1,0.05}) = 0.90$$
$$P(\chi^2_{n-1,0.95} \leq \frac{(n-1)\hat{\sigma}^2}{\sigma^2} \leq \chi^2_{n-1,0.05}) = 0.90$$
$$P\left(\frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1,0.05}} \leq \sigma^2 \leq \frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1,0.95}}\right) = 0.90$$

Then, the 0.9 confidence interval for σ^2 is

$$\left[\frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1,0.05}}, \frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1,0.95}}\right] \tag{2}$$

Numerical values:

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{9}\sum_{i=1}^{10} X_i^2 - \frac{10}{9}\left[\frac{1}{10}\sum_{i=1}^{10} X_i\right]^2 = \frac{45}{9} - \frac{10}{9}\left(\frac{21}{10}\right)^2 = 0.1 \\
\chi^2_{9,0.05} &= 16.919 \\
\chi^2_{9,0.95} &= 3.325
\end{aligned}$$

0.9 confidence interval for σ^2: $[0.053, 0.271]$

d) Lets use $\hat{\sigma}^2 = s$ to leave the notation uncluttered

$$V = \frac{(n-1)S}{\sigma^2}, \quad S = \frac{\sigma^2 V}{(n-1)}$$

S is a function of V, which has density

$$f_v(v) = \frac{1}{2^{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})} v^{\frac{n-1}{2}-1} e^{-v/2}, v \geq 0.$$

S is a one-to-one transformation of a continuous random variable V, then we need to apply the following formula to compute the density of S,

$$\begin{aligned}
f_s(s) &= f_v\left(\frac{(n-1)s}{\sigma^2}\right)\left|\frac{dv}{ds}\right| \\
&= \frac{1}{2^{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})}\left[\frac{(n-1)s}{\sigma^2}\right]^{\frac{n-1}{2}-1} \exp\left\{-\frac{(n-1)s}{\sigma^2 2}\right\}\frac{(n-1)}{\sigma^2} \\
&= \frac{1}{[\frac{2\sigma^2}{n-1}]^{(n-1)/2}\Gamma((n-1)/2)} s^{(n-1)/2-1} \exp\left\{-\frac{(n-1)s}{2\sigma^2}\right\}
\end{aligned}$$

The equation above has the form of a Gamma distribution. We can write that

$$
\begin{aligned}
f_s(s) &= \frac{1}{[\frac{2\sigma^2}{n-1}]^{(n-1)/2}\Gamma((n-1)/2)} s^{(n-1)/2-1} \exp\left\{-\frac{(n-1)s}{2\sigma^2}\right\} \\
&= \frac{1}{\beta^\alpha \Gamma(\alpha)} s^{\alpha-1} \exp\{-s/\beta\}, \quad s \geq 0
\end{aligned}
$$

with $\beta = \frac{2\sigma^2}{n-1}, \quad \alpha = \frac{n-1}{2}$.
Then, the expectation and the variance of S can be computed straitforwardly as
$E(S) = E(\hat{\sigma}^2) = \alpha\beta = \frac{n-1}{2}\frac{2\sigma^2}{n-1} = \sigma^2$, unibiased.
$Var(S) = Var(\hat{\sigma}^2) = \alpha\beta^2 = \frac{n-1}{2}\frac{4\sigma^4}{(n-1)^2} = \frac{2\sigma^4}{(n-1)}$.

Oppgave 3

$X_1, ..., X_n$ i.i.d. $f(x;\beta) = \frac{\beta 2^\beta}{x^{\beta+1}}$, $2 \leq x \leq \infty$

a) Since $X_1, ..., X_n$ are independent, the likelihood function is given by

$$
L(\beta; x_1, ..., x_n) = \prod_{i=1}^{n} \frac{\beta 2^\beta}{x_i^{\beta+1}} = \beta^n 2^{n\beta} \prod_{i=1}^{n} \frac{1}{x_i^{\beta+1}}
$$

To find the maximum of the likelihood function, it is easier to maximize the log-likelihood, given by

$$
\ln L(\beta; x_1, ..., x_n) = n \ln\beta + n\beta \ln 2 - (\beta+1) \sum_{i=1}^{n} \ln x_i
$$

We then set the first derivative of the log-likelihood equals to zero:

$$
\frac{d \ln L(\cdot)}{d\beta} = \frac{n}{\beta} + n \ln 2 - \sum_i \ln x_i = 0
$$

We then have

$$\hat{\beta} = \frac{n}{\sum_i \ln x_i - n \ln 2} = \frac{n}{\sum_{i=1}^{n} \ln(x_i/2)}$$

The maximum likelihood estimator is:

$$\hat{\beta} = \frac{n}{\sum_{i=1}^{n} \ln(X_i/2)}$$