University of Stavanger

EXAM IN: STA500 INTRODUCTION TO PROBABILITY AND STATISTICS 2

DURATION: 4 HOURS DATE: DECEMBER 5, 2015

PERMITTED AIDS: Approved simple calculator (HP30S, Casio FX82, TI-30, Citizen SR-270X, Texas BA II Plus or HP17bII+).

THE EXAM CONSISTS OF 3 PROBLEMS ON 2 PAGES, 9 PAGES OF ENCLOSURES.

COURSE RESPONSIBLE: Tore Selland Kleppe PHONE: 51 83 17 17

Problem 1: A marine researcher wishes to determine the distribution of the length of a particular species of fish in the North Sea. It is assumed that the distribution of the length of a fish X is exponential with mean β, i.e.

$$f(x;\beta) = \frac{1}{\beta}\exp\left(-\frac{x}{\beta}\right).$$

The researcher receives n such fish from a commercial trawler. The equipment of the trawler is made so that only fish with length greater than $c > 0$ gets caught by the trawl.

a) Show that the distribution of the length Y of a fish provided to the researcher from the trawler will have the density

$$f(y;\beta) = \frac{1}{\beta}\exp\left(-\frac{y-c}{\beta}\right),\ y > c.$$

b) Show that $E(Y) = \beta + c$ and $Var(Y) = \beta^2$.

To estimate the population mean parameter β, the researcher first consider using the estimator

$$\tilde{\beta} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

where $Y_i, i = 1,\dots,n$ are independent lengths of fish received from the trawler.

c) Find the mean and variance of $\tilde{\beta}$.
Is $\tilde{\beta}$ an unbiased estimator for β?
Is $\tilde{\beta}$ a consistent estimator for β?

d) Based on the above expression for $f(y;\beta)$, find the maximum likelihood estimator $\hat{\beta}$ for β.

e) Show that the maximum likelihood estimator $\hat{\beta}$ is consistent.
Find a 95% Wald confidence interval for β based on the maximum likelihood estimator in d).

Problem 2: Consider the Markov chain model $\{X_t,\ t = 0, 1, \dots\}$ with state space $\mathcal{S} = [0, 1]$ and transition probability matrix

$$P = \begin{bmatrix} (1-\lambda) & \lambda \\ \lambda & (1-\lambda) \end{bmatrix},\ 0 < \lambda < 1.$$

a) State the requirements for this Markov chain to have steady state probabilities. Show that the steady state probabilities are $\pi_0 = \pi_1 = 1/2$.
Draw a transition graph for the Markov chain.

The quantity

$$\rho = \frac{E\left[(X_t - E(X_t))(X_{t+1} - E(X_{t+1}))\right]}{Var(X_t)}$$

is known as the first order autocorrelation of the Markov chain X_t. I.e. it is the correlation between X_t and X_{t+1}. Moreover, the joint probability mass function of (X_t, X_{t+1}) is given as

$$P(X_t = i, X_{t+1} = j) = \pi_i p_{ij},\ i, j = 0, 1.$$

b) Find $E(X_t)$ and $E(X_{t+1})$.
Find $Var(X_t)$.
Compute the first order autocorrelation ρ for the process X_t.
Give an interpretation of how the parameter λ influences the behavior of the chain.

Problem 3: To model the life time X of a particular electronic component, an engineer uses a log-normal distribution with precision parameter τ. This distribution has probability density function given by

$$f(x;\tau) = \sqrt{\frac{\tau}{2\pi}}\frac{1}{x}\exp\left(-\frac{1}{2}\tau(\log x)^2\right),\ x, \tau > 0.$$

The engineer takes a Bayesian approach and uses a gamma(α, β) prior. Suppose the engineer has access to life time data $X_1, \dots, X_n \sim$ iid $f(x;\tau)$.

a) Show that the posterior distribution for τ, i.e. $p(\tau|X_1, \dots, X_n)$, is a gamma(α^*, β^*) distribution with shape and scale parameters

$$\alpha^* = n/2 + \alpha,\ \beta^* = \left(\frac{1}{2}\sum_{i=1}^{n}(\log X_i)^2 + \frac{1}{\beta}\right)^{-1}.$$

The engineer has $n = 4$ observations: 1.1, 2.0, 0.4, 0.3 and the prior parameters are selected to be $\alpha = 10$ and $\beta = 0.1$.

b) Based on the prior and data, find the Bayes estimator $\hat{\tau}_{Bayes} = E(\tau|X_1, \dots, X_n)$.
Based on the prior and data, find a 95% credible interval for τ.

Solutions

1.a
The researcher receives censored samples Y from the population X with distribution of Y being that of $X|X > c$. Now

$$f(y) = \frac{f_X(y)}{P(X > c)} = \frac{\frac{1}{\beta}\exp\left(-\frac{y}{\beta}\right)}{1 - F(c)} = \frac{\frac{1}{\beta}\exp\left(-\frac{y}{\beta}\right)}{\exp\left(-\frac{c}{\beta}\right)} = \frac{1}{\lambda}\exp\left(-\frac{y-c}{\beta}\right).$$

1.b
Use e.g. integral formulas in tables and formulas

$$E(Y) = \int_c^\infty yf(y)dy = c + \beta.$$

Moreover

$$E(Y^2) = \int_c^\infty y^2 f(y)dy = 2\beta^2 + 2c\beta + c^2,$$

and therefore

$$Var(Y) = E(Y^2) - E(Y)^2 = 2\beta^2 + 2c\beta + c^2 - c^2 - 2\beta c - \beta^2 = \beta^2.$$

1.c
Based on 1.b we have that

$$E(\tilde{\beta}) = \frac{1}{n}\sum_i E(Y_i) = \frac{1}{n}n(c + \beta) = c + \beta.$$

$$Var(\tilde{\beta}) = \frac{1}{n^2}\sum_i Var(Y_i) = \frac{1}{n^2}n\beta^2 = \frac{\beta^2}{n}.$$

Note that $E(\tilde{\beta}) = c + \beta \neq \beta$ and therefore the estimator is biased (i.e. not unbiased). Moreover, the bias does not vanish as n grow, and therefore the estimator is not consistent.
1.d
Likelihood function

$$L(\beta) = \prod_{i=1}^n f(y_i; \beta) = \beta^{-n}\exp\left(-\frac{1}{\beta}\sum_i (y_i - c)\right),$$

log-likelihood function

$$l(\beta) = -n\log(\beta) - \frac{1}{\beta}\sum_i (y_i - c),$$

First derivative wrt β:

$$\frac{\partial}{\partial\beta}l(\beta) = -\frac{n}{\beta} + \frac{1}{\beta^2}\sum_i (y_i - c)$$

Solve for critical point:

$$\begin{aligned}
0 &= -\frac{n}{\beta} + \frac{1}{\beta^2}\sum_i (y_i - c) \\
&\Downarrow \\
n\beta &= \sum_i (y_i - c) \\
&\Downarrow \\
\hat{\beta} &= \frac{1}{n}\sum_i (y_i - c) = \bar{y} - c
\end{aligned}$$

Check that this is a maximizer:

$$\frac{\partial^2}{\partial \beta^2} l(\hat{\beta}) = \frac{n}{\hat{\beta}^2} - \frac{2\overbrace{\sum_i (y_i - c)}^{n\hat{\beta}}}{\hat{\beta}^3} = -\frac{n}{\hat{\beta}^2} < 0.$$

I.e. $\hat{\beta}$ correspond to a maximum of the log-likelihood function.
1.e
The estimator is consistent as it is unbiased

$$E(\hat{\beta}) = \underbrace{E(\bar{y})}_{=\beta+c} - c = \beta$$

and the variance vanishes as $n \to \infty$:

$$Var(\hat{\beta}) = Var(\bar{y}) = \beta^2/n \to 0.$$

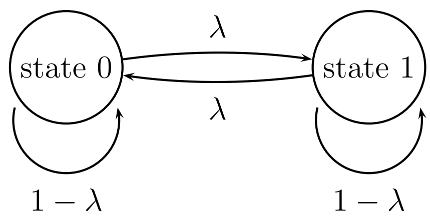
Wald-type 95% confidence interval (found second derivative above):

$$\left[\hat{\beta} \mp 1.96\frac{\hat{\beta}}{\sqrt{n}}\right].$$

2.a
The finite state space chain is irreducible (the two states communicate) and aperiodic (as e.g. $p_{00} > 0$), and therefore admit steady state probabilities. These are found by solving e.g.

$$\begin{aligned}
\pi_0 &= (1-\lambda)\pi_0 + \lambda\pi_1, \\
1 &= \pi_0 + \pi_1. \\
&\Downarrow \\
\pi_0 &= (1-\lambda)\pi_0 + \lambda(1-\pi_0) \\
&\Downarrow \\
2\lambda\pi_0 &= \lambda \\
&\Downarrow \\
\pi_0 &= \frac{1}{2} \\
&\Downarrow \\
\pi_1 &= 1 - \pi_0 = \frac{1}{2}.
\end{aligned}$$

I.e. in the long run, the chain spend equal amount of time in both states. The transition graph is drawn below:



2.b
Expectations:

$$E(X_t) = \sum_{i,j=0,1} i\pi_i p_{ij} = \sum_i i\pi_i \underbrace{\sum_j p_{ij}}_{=1} = 0 \cdot 1/2 + 1 \cdot 1/2 = 1/2.$$

$$E(X_{t+1}) = \sum_{i,j=0,1} j\pi_i p_{ij} = \underbrace{\pi_0 p_{00} \cdot 0 + \pi_0 p_{01} \cdot 1}_{i=0} + \underbrace{\pi_1 p_{10} \cdot 0 + \pi_1 p_{11} \cdot 1}_{i=1} = 1/2 \cdot \lambda + 1/2 \cdot (1-\lambda) = 1/2.$$

Alternatively, reasoning from the fact that the steady state probabilities are the marginals of both X_t and X_{t+1} is also OK.
The variance:

$$Var(X_t) = \pi_0(0-1/2)^2 + \pi_1(1-1/2)^2 = 1/4.$$

Alternatively, going the trough the joint distribution as above is also OK.
The first order autocorrelation can then be completed as

$$\begin{aligned} E\left[(X_t - E(X_t))(X_{t+1} - E(X_{t+1}))\right] &= \sum_{i,j=0,1} (i-1/2)(j-1/2)\pi_i p_{ij} \\ &= \underbrace{(-1/2)(-1/2)1/2(1-\lambda) + (-1/2)(1/2)1/2\lambda}_{i=0} \\ &\quad + \underbrace{(1/2)(-1/2)1/2\lambda + (1/2)(1/2)1/2(1-\lambda)}_{i=1} \\ &= 1/8(1-\lambda) - 1/8\lambda - 1/8\lambda + 1/8(1-\lambda) \\ &= 1/4 - \lambda/2. \end{aligned}$$

Therefore

$$\rho = \frac{1/4 - \lambda/2}{1/4} = 1 - 2\lambda.$$

The parameter λ controls the dependence structure of the chain, without altering the steady state distribution. I.e. for small λ, i.e. $\lambda < 1/2$, X_{t+1} tends to be equal to X_t. For $\lambda = 1/2$, the process has no autocorrelation (X_{t+1} is independent of X_t in this case). For $\lambda > 1/2$ the process tends to switch state more often than it remains in the state.
3.a
Likelihood:

$$L(\tau) \propto \tau^{n/2} \exp\left(-\frac{1}{2}\tau \sum_{i=1}^{n} (\log X_i)^2\right)$$

Prior:

$$p(\tau) \propto \tau^{\alpha-1} \exp(-\tau/\beta)$$

Posterior:

$$
\begin{aligned}
p(\tau|X_1,\ldots,X_n) &\propto \tau^{n/2}\tau^{\alpha-1}\exp\left(-\frac{1}{2}\tau\sum_{i=1}^{n}(\log X_i)^2\right)\exp(-\tau/\beta) \\
&= \tau^{n/2+\alpha-1}\exp\left(-\tau\left(\frac{1}{2}\tau\sum_{i=1}^{n}(\log X_i)^2+\frac{1}{\beta}\right)\right).
\end{aligned}
$$

We recognize the posterior kernel to be a gamma(α^*,β^*) distribution with shape parameter

$$\alpha^* = n/2+\alpha$$

and scale parameter

$$\beta^* = \left(\frac{1}{2}\sum_{i=1}^{n}(\log X_i)^2+\frac{1}{\beta}\right)^{-1}.$$

3.b

First we compute that $\sum_{i=1}^{n}(\log x_i)^2 = 2.778676$. The mean under a gamma distribution is $\alpha\beta$, and therefore

$$\hat{\tau}_{Bayes} = \alpha^*\beta^* = \frac{4/2+10}{\frac{2.778676}{2}+10} = 1.053617.$$

Given that $\tau|X_1,\ldots,X_n \sim$ gamma(α^*,β^*), we use the relation between general gamma distributions and χ^2-distributions to arrive at

$$
\begin{aligned}
P\left(\chi^2_{1-\alpha/2,2\alpha^*} < \frac{2\tau}{\beta^*} < \chi^2_{\alpha/2,2\alpha^*}|X_1,\ldots,X_n\right) &= 1-\alpha \\
&\Downarrow \\
P\left(\frac{\beta^*}{2}\chi^2_{1-\alpha/2,2\alpha^*} < \tau < \frac{\beta^*}{2}\chi^2_{\alpha/2,2\alpha^*}|X_1,\ldots,X_n\right) &= 1-\alpha
\end{aligned}
$$

Now, in our case $\alpha^* = 12$ and therefore $\chi^2_{0.975,2\cdot 12} = 12.401$, $\chi^2_{0.025,2\cdot 12} = 39.364$. Moreover $\beta^* = 0.08780141$ and thus

$$\tau_L = 0.5\cdot 0.08780141\cdot 12.401 = 0.54,\ \tau_U = 0.5\cdot 0.08780141\cdot 39.364 = 1.73$$

where $[\tau_L,\tau_U]$ defines the sought credible interval.