University of Stavanger

EXAM IN: STA500 INTRODUCTION TO PROBABILITY AND STATISTICS 2

DURATION: 4 HOURS DATE: DECEMBER 9th, 2017

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

THE EXAM CONSISTS OF 5 PROBLEMS ON 4 PAGES, 18 PAGES OF ENCLOSURES.

COURSE RESPONSIBLE: Tore Selland Kleppe PHONE: 95 92 23 64

Note: Throughout this exam, all logarithms are natural logarithms, so that $\log(x) = \ln(x)$, and 10-based logarithms are not used.

Problem 1: The life time, measured in years, T, of an electronic component has a Weibull distribution with parameters $\beta = 1.5$ and $\alpha = 0.02$. The corresponding probability density is given as

$$f_T(t) = 0.03\sqrt{t}\exp(-0.02t^{1.5}),$$

and the cumulative distribution function is

$$F_T(t) = 1 - \exp(-0.02t^{1.5})$$

a) Find the probability that a component of this type will have a life time of more than 10 years.
Find the probability that a component of this type will have a life time between 10 and 15 years.
Given that the component is still working after 10 years, what is the probaility that it will have a life time of less than 15 years.

In an electronic device, 12 such electronic components are installed. Assume that these 12 components fail independently of each other. Let Y be the number of components still working after 10 years.

b) What is the probability distribution of Y? Explain.
Calculate $E(Y)$, $Var(Y)$ and $P(Y = 0)$.

Now, consider another electronic device that consists of two components of the type discussed above, where the life times of these two components are independent. For this device to work, both components need to be working. Thus, the life time of the device is

$$U = \min(T_1, T_2),$$

where T_1, T_2 are distributed as T above.

c) Find the distribution of the life time of the device, U.
Calculate $P(U > 10)$ and compare your answer to $P(T > 10)$ above.

Problem 2: Consider observations $X_1, \ldots, X_n$ that are iid distributed with probability density

$$f_{X_i}(x) = \exp\left(-\frac{1}{2}x^2 \exp(\lambda) + \frac{\lambda}{2}\right) \frac{1}{\sqrt{2\pi}},$$

i.e. X_i has a normal distribution with mean 0 and variance $\exp(-\lambda)$.

a) Write down the likelihood function and find the maximum likelihood estimator $\hat{\lambda}$ of λ.

A random variable Y is said to have an exp-gamma(a, b)-distribution if the probability density of Y is

$$f_Y(y) = \frac{1}{\Gamma(a)b^a} \exp\left(ya - \frac{\exp(y)}{b}\right), a > 0, b > 0, -\infty < y < \infty.$$

Namely, if $Y \sim$ exp-gamma(a, b), then $\exp(Y) \sim$ gamma(a, b).

b) Suppose you choose an exp-gamma(a_0, b_0) prior for λ. Find the posterior distribution of λ based on this prior and observations $X_1, \ldots, X_n$.

An alternative to the regular Bayes estimator (posterior mean) is the MAP[1] $\tilde{\lambda}$, which defined in this case as the value of λ that maximizes the posterior density, i.e.

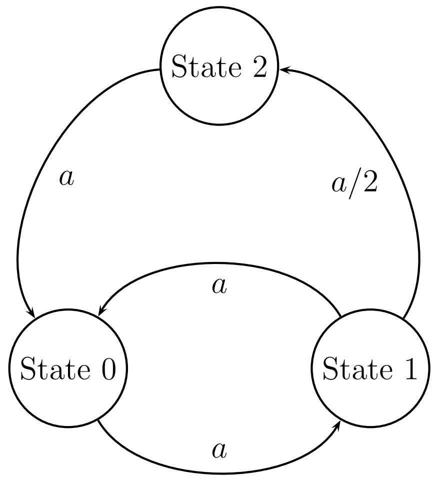
$$\tilde{\lambda} = \arg\max_{\lambda} p(\lambda | T_1, \ldots, T_n).$$

c) Find the MAP $\tilde{\lambda}$ for λ in the situation considered in points a-b). Compare $\tilde{\lambda}$ to the MLE $\hat{\lambda}$.

d) Find a $(1-\alpha) \times 100\%$ Bayesian credible interval for λ. Hint: use the distribution of $Y = \exp(X)$ when X has an exp-gamma distribution.

[1] Abbreviation for maximum a-posteori (density).

Problem 3: Consider at continuous time Markov chain $X(t), t \geq 0$ with state space $\{0, 1, 2\}$ and transition graph:



where $a > 0$. Note that the indicated numbers are specific transition rates.

a) Suppose first that $X(0) = 0$. Find the mean and variance of the time until the process first leaves state 0.
What is the probability that the process will jump to state 2 when it leaves state 1?
Why does this process admit steady state probabilities?
Is this process a birth-and-death process?

b) Find the steady state probabilities π_0, π_1, π_2.
Comment on why or why not the steady state probabilities depend on a, and give an interpretation of this fact.

Problem 4: Consider the (discrete time) Markov chain, specified in terms of the transition matrix:

$$P = \begin{pmatrix} 4/5 & 0 & 0 & 1/10 & 1/10 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

a) Draw the transition graph and find the classes of the Markov chain.
Is the process irreducible?
Are the classes transient or recurrent?
What is the period of each of the classes?

Problem 5: Let $Z_t \sim$ iid Bernoulli(p), $0 < p < 1$, that is, Z_t is a Bernoulli process with $P(Z_t = 1) = p$ and $P(Z_t = 0) = 1 - p$ independently over time. Moreover, define the stochastic process X_t so that

$$X_t = \begin{cases} 0 & \text{if } Z_t = 0,\ Z_{t-1} = 0 \\ 1 & \text{if } Z_t = 1,\ Z_{t-1} = 0 \\ 2 & \text{if } Z_t = 0,\ Z_{t-1} = 1 \\ 3 & \text{if } Z_t = 1,\ Z_{t-1} = 1 \end{cases}$$

a) Argue for why the process X_t is a Markov chain and show that the transition probability matrix for X_t is given by

$$P = \begin{pmatrix} 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p \\ 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p \end{pmatrix}.$$

b) Why does X_t admit steady state probabilities?
Argue for why the steady state probabilities are $\pi_i = p^2_{ji}$ for all $i, j \in \{0, 1, 2, 3\}$, and find the steady state probabilities.

Solutions

1,a)
First question: $P(X > 10) = 1 - P(X < 10) = 1 - F_T(10) = 0.5312856$.
Second question: $P(10 < X < 15) = P(X < 15) - P(X < 10) = F_T(15) - F_T(10) = 0.2183929$.
Third question: $P(X < 15 | X > 10) = P(10 < X < 15)/P(X > 10) = 0.4659403$.

1,b)
Y is the sum of $n = 12$ independent $0-1$ outcome trials, where an 1-outcome correspond to the event that a component is still working after 10 years. Thus Y has a binomial$(12, 0.5312856)$-distribution.
Expectation and variance: $E(Y) = np = 12 \times 0.5312856 = 6.375427$, $Var(Y) = np(1-p) = 2.988255$.
Probability of all failing: $P(Y = 0) = (1-p)^n = 0.0001124345$.

1,c)
The CDF of U is

$$F_U(u) = 1 - (1 - F_T(u))^2 = 1 - \exp(-0.02t^{1.5})^2 = 1 - \exp(-0.04t^{1.5})$$

Namely, U also has a Weibull distribution with shape parameter $\beta = 1.5$ and $\alpha = 0.04$.
The sought probaility is $P(U > 10) = 1 - F_U(10) = 0.282264398$.
It is seen that (since this a series system) that the device involving two components typically have shorter life times than each individual component.

2,a)
Likelihood function:

$$L(\lambda; X_1, \ldots, X_n) = \prod_{i=1}^{n} f_{X_i}(X_i) \propto \exp\left(\frac{n}{2}\lambda - \frac{\exp(\lambda)}{2}\sum_{i=1}^{n} X_i^2\right).$$

MLE is found as maximizer of $\log(L(\lambda))$:

$$\begin{aligned}
\frac{\partial}{\partial\lambda}\log(L(\lambda)) &= \frac{n}{2} - \frac{\exp(\lambda)}{2}\sum_{i=1}^{n} X_i^2 = 0 \\
&\Downarrow \\
n &= \exp(\lambda)\sum_{i=1}^{n} X_i^2 \\
&\Downarrow \\
\hat{\lambda} &= -\log\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right)
\end{aligned}$$

2,b)

Posterior kernel is

$$\exp\left(\frac{n}{2}\lambda - \frac{\exp(\lambda)}{2}\sum_{i=1}^{n} X_i^2\right)\exp(\lambda a_0 - \exp(\lambda)/b_0)$$
$$= \exp\left(\lambda\left(\frac{n}{2}+a_0\right) - \exp(\lambda)\left(\frac{1}{2}\sum_{i=1}^{n} X_i^2 + \frac{1}{b_0}\right)\right)$$

Thus, the posterior is exp-gamma with parameters $a = a_0 + n/2$ and

$$b = \frac{1}{\frac{1}{2}\sum_{i=1}^{n} X_i^2 + \frac{1}{b_0}}$$

2,c)
For a general exp-gamma-distribution (in λ), we have that

$$\frac{\partial}{\partial\lambda}\log f(\lambda) = a - \exp(\lambda)/b \Rightarrow \tilde{\lambda} = \log(ab).$$

Thus, in the case considered here,

$$\tilde{\lambda} = \log\left(\frac{n/2 + a_0}{\frac{1}{2}\sum_{i=1}^{n} X_i^2 + \frac{1}{b_0}}\right) = -\log\left(\frac{\sum_{i=1}^{n} X_i^2 + \frac{2}{b_0}}{n + 2a_0}\right).$$

Notice that the MAP and MLE are asymptotically (as $n \to \infty$) the same, as both $\sum_i X^2$ and n go to infinity at the same rate. Moreover, the Bayesian treatment of this model with given prior effectively correspond to adding $2a_0$ observations with sum of squares $2/b_0$ to the sample.

2,d)
Using the hint, we have that

$$\lambda | T_1, \ldots, T_n \sim \text{exp-gamma}(a, b),$$

implies

$$\exp(\lambda) | T_1, \ldots, T_n \sim \text{gamma}(a, b).$$

Using the regular transformation (from gamma to χ_2), we have that

$$\frac{2\exp(\lambda)}{b} | T_1, \ldots, T_n \sim \chi^2_{2a}.$$

Thus

$$\begin{aligned}
1-\alpha &= P\left(\chi^2_{1-\alpha/2,2a} < \frac{2\exp(\lambda)}{b} < \chi^2_{\alpha/2,2a} | T_1, \ldots, T_n\right) \\
&= P\left(\chi^2_{1-\alpha/2,2a}\frac{b}{2} < \exp(\lambda) < \frac{b}{2}\chi^2_{\alpha/2,2a} | T_1, \ldots, T_n\right) \\
&= P\left(\log\left(\chi^2_{1-\alpha/2,2a}\frac{b}{2}\right) < \lambda < \log\left(\chi^2_{\alpha/2,2a}\frac{b}{2}\right) | T_1, \ldots, T_n\right)
\end{aligned}$$

Plugging in the values for a, b, we obtain the credible interval

$$\left[\log\left(\frac{\chi^2_{1-\alpha/2,2a_0+n}}{\sum_{i=1}^n X_i^2 + \frac{2}{b_0}}\right), \log\left(\frac{\chi^2_{\alpha/2,2a_0+n}}{\sum_{i=1}^n X_i^2 + \frac{2}{b_0}}\right)\right]$$

3,a)
The time until the process first leaves state 0 is exponentially distributed with mean $1/a$ and variance $1/(a^2)$.
Probability of jump $1 \to 2$ given that a jump out of 1 occurs: $p_{12} = q_{12}/\nu_1 = 0.5a/(1.5a) = 1/3 = 0.333333$.
Since the process has a finite state space, it is sufficient for the process to be irreducible for steady state probabilities to exist. The process is indeed irreducible as all states are reachable from any other state (in at most 2 transitions).
The process is not a birth and death process since it is possible to jump from state 2 to state 0, and circumventing state 1.

3,b)
We write down the balance equations for states 1,2 (as these have fewest incoming arcs):

$$\begin{aligned} \text{state 1:} \quad 0 = &\ a\pi_0 - \frac{3a}{2}\pi_1, \\ \text{state 2:} \quad 0 = &\ \frac{a}{2}\pi_1 - a\pi_2, \\ \text{normalization:} \quad 1 = &\ \pi_0 + \pi_1 + \pi_2. \end{aligned}$$

We see from the two former equations that π_0, π_2 are easily expressed in terms of π_1, i.e.:

$$\pi_0 = \frac{3}{2}\pi_1, \quad \pi_2 = \frac{1}{2}\pi_1.$$

Plugging these two into the normalization relation, we have that

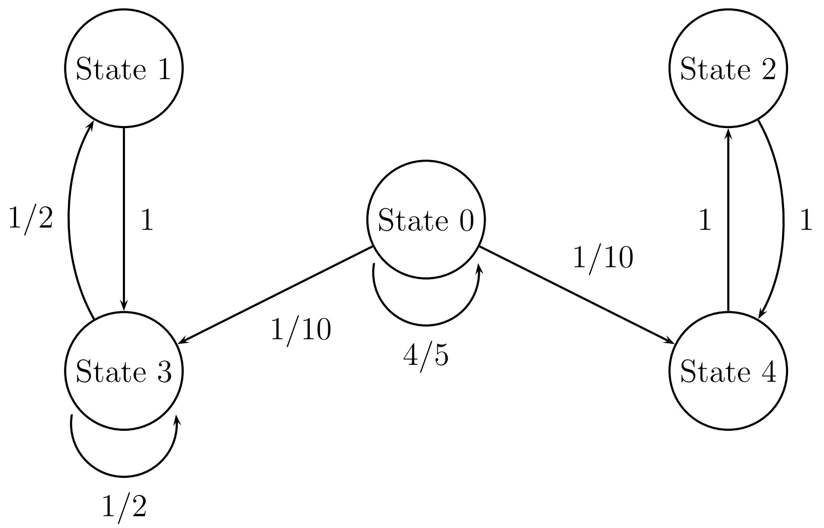
$$1 = \frac{3}{2}\pi_1 + \pi_1 + \frac{1}{2}\pi_1 = 3\pi_1 \Rightarrow \pi_1 = \frac{1}{3}.$$

$$\Rightarrow \pi_0 = \frac{3}{2}\pi_1 = \frac{1}{2}.$$

$$\Rightarrow \pi_2 = \frac{1}{2}\pi_1 = \frac{1}{6}.$$

Note that the constant a does not influence the steady state probabilities. That is because it enters multiplicatively in all of the specific transition rates, and effectively modifies the time unit, but does nothing to the over all dynamics once rescaled time is taken into account.

4,a)
The transition graph is something like



The process has three classes, and is therefore not irreducible.

- Class $\{0\}$ transient.
- Class $\{1,3\}$ recurrent, aperiodic
- Class $\{2,4\}$ recurrent, period 2

5,a)
The process X_t is a Markov chain as there is no additional information about X_t (depending on Z_t, Z_{t-1} in knowing X_{t-2} (depending on Z_{t-2}, Z_{t-3}) or any other $X_{t-k},\ k = 2, \ldots$.

The process admit the following transitions

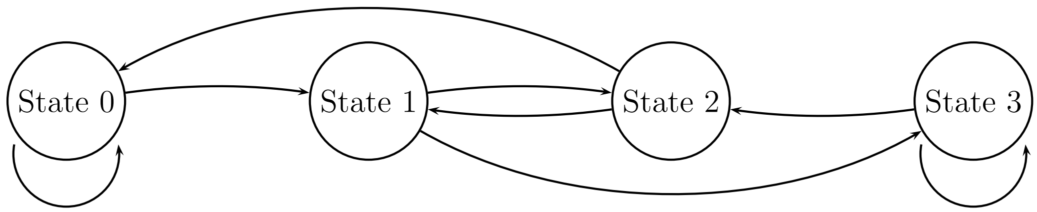
- $X_{t-1} = 0 \to Z_{t-1} = 0:$ Transitions to 0 ($Z_t = 0$, prob=$1-p$) or 1 ($Z_t = 1$, prob=p) possible
- $X_{t-1} = 1 \to Z_{t-1} = 1:$ Transitions to 2 ($Z_t = 0$, prob=$1-p$) or 3 ($Z_t = 1$, prob=p) possible.
- $X_{t-1} = 2 \to Z_{t-1} = 0:$ Transitions to 0 ($Z_t = 0$, prob=$1-p$) or 1 ($Z_t = 1$, prob=p) possible.
- $X_{t-1} = 3 \to Z_{t-1} = 1:$ Transitions to 2 ($Z_t = 0$, prob=$1-p$) or 3 ($Z_t = 1$, prob=p) possible.

This gives us the transition probability matrix

$$P = \begin{pmatrix} 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p \\ 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p \end{pmatrix}.$$

5,b)

In order to determine whether X_t is irreducible and aperiodic, it is helpful to draw the transition graph of the process (for simplicity, skipping labeling the arcs):



The process is irreducible (all states are reachable from every other in two transitions), and the single class is aperiodic as e.g. $p_{00} > 0$.

Looking carefully at the definition of the process, we see that X_t (which depend on Z_t, Z_{t-1}) and X_{t+2} (which depend on Z_{t+2}, Z_{t+1}) are independent in this case, and therefore

$$P(X_{t+2} = i | X_t = j) = P(X_{t+2} = i) = \pi_i.$$

Thus the steady state probabilities are easily found by calculating any row of P^2, e.g. here we compute the first row:

$$\pi_0 = p_{00}^2 = (1-p)^2 + p \times 0 + 0 \times (1-p) + 0 \times 0 = (1-p)^2,$$

and similiarly:

$$\pi_1 = p_{01}^2 = p(1-p), \ \pi_2 = p_{02}^2 = p(1-p), \ \pi_3 = p_{03}^2 = p^2.$$

Alternatively, one could have found the steady state probabilities by directly reasoning from the definition of X_t, e.g. $\pi_0 = P(X_t = 0) = P(Z_t = 0, Z_{t-1} = 0) = P(Z_t = 0)P(Z_{t-1} = 0) = (1-p)^2$ and so on.