University of Stavanger

EXAM IN: STA500 INTRODUCTION TO PROBABILITY AND STATISTICS 2

DURATION: 4 HOURS DATE: NOVEMBER 18, 2013

PERMITTED AIDS: Approved simple calculator (HP30S, Casio FX82, TI-30 or Citizen SR-270X).

One yellow A4 size sheet with your own handwritten notes.

THE EXAM CONSISTS OF 3 PROBLEMS ON 4 PAGES, 9 PAGES OF ENCLOSURES.

Problem 1:

For concrete modules of a certain kind it is known that the compressive strength (measured in how many mega-pascal pressure the module can withstand before it is crushed) has a normal distribution with expectation $\mu = 49$ and standard deviation $\sigma = 3$. Assume that the compressive strengths for different modules are independent.

a) Show that the probability that a module has a compressive strength of more than 45 is 0.91 (rounded to two decimal places).
What is the probability that a module has a compressive strength of between 45 and 50?
If we have 10 modules of this kind, what is the probability that the weakest of them has a compressive strength of more than 45?

We need 10 modules of the considered kind in a construction. We require that all modules we shall use must have a compressive strength of more than 45. We will thus test modules until we have found 10 modules that have a compressive strength of more than 45. Let Y be the number of modules we have to test until we have found 10 modules that have a compressive strength of more than 45.

b) Which probability distribution does Y have? Explain why.
Calculate $P(Y \geq 12)$.
Find $\mathrm{E}(Y)$.

Problem 2:

Cars wanting to use the automatic car wash machine at a gas station arrive as a Poisson process with rate λ per hour. A car wash takes 6 minutes.

We shall in this problem consider a situation where $\lambda = 8$ per hour.

a) What is the probability that at least one new car arrive to the car wash during the 6 minutes a car is being washed?
What is the probability that the time between two consecutive car arrivals to the car wash is less than 10 minutes?
From the start of a car wash, what is the probability that the time until two new cars have arrived to the car wash is less than 6 minutes?

If new cars arrive while a car is being washed they wait in line, but due to space restrictions the gas station only allow until three cars to wait in line. If more cars arrive for a car wash while three cars are waiting in line they are directed to a different car wash in the neighbourhood.

Let X_n be the number of cars waiting in line when car number n is finished with the wash (just before the first car in line can enter the wash). Then the process $\{X_n :\ n = 0, 1, 2, 3, \ldots\}$ is a Markov chain with transition probability matrix:

$$P = \begin{pmatrix} 0.45 & 0.36 & 0.14 & 0.05 \\ 0.45 & 0.36 & 0.14 & 0.05 \\ 0 & 0.45 & 0.36 & 0.19 \\ 0 & 0 & 0.45 & 0.55 \end{pmatrix}$$

b) Show how we get the numbers in the third row of the transition probability matrix in this situation. (I.e. explain how we get the given numbers for the matrix elements p_{20}, p_{21}, p_{22} and p_{23}.)
If three cars are waiting in line when a car wash is finished, what is the expected number of car washes we will see until there is less than three cars waiting in line when a wash is finished?

It is given that:

$$P^2 = \begin{pmatrix} 0.36 & 0.36 & 0.19 & 0.09 \\ 0.36 & 0.36 & 0.19 & 0.09 \\ 0.20 & 0.32 & 0.28 & 0.20 \\ 0 & 0.20 & 0.41 & 0.39 \end{pmatrix}$$

c) Find:
$P(X_2 = 2|X_1 = 0)$, $P(X_{12} = 2|X_{10} = 1)$, $P(X_5 = 3|X_3 = 2, X_2 = 1, X_1 = 1)$, $P(X_{n+2} = 1, X_{n+1} = 2|X_n = 3)$ and $P(X_6 = 1, X_5 = 2, X_4 = 3|X_2 = 2, X_1 = 1)$.

It can be shown that the steady state probabilities of the Markov chain are $\pi_0 = 0.26$, $\pi_1 = 0.32$, $\pi_2 = 0.25$ and $\pi_3 = 0.17$.

d) Set up the equations we need to solve to find the steady state probabilities (do not try to solve the equations, the solution is given above).
What is in the long run the expected number of cars waiting in line when a car wash is finished?
If there are three cars in line when a wash is finished, what is the expected number of cars in line when the next wash is finished? Compare the two expectations and comment briefly.
How would the steady state probabilities and the expectations change if we instead of $\lambda = 8$ had a situation with an arrival rate of $\lambda = 12$ per hour? Without doing any calculations, argue in which direction (increase or decrease) π_0, π_3 and the two expectations calculated above would change if $\lambda = 12$.

Problem 3:

The lifetime of a certain type of detectors used on offshore installations is assumed to have an exponential distribution with expectation β. Data on lifetimes of such detectors used on different installations are found in two different reliability data bases. Let $T_1, T_2, \ldots, T_n$ denote lifetimes reported to the first data base, and let $S_1, S_2, \ldots, S_m$ denote lifetimes reported to the second data base. We assume that all these lifetimes are independent and exponentially distributed with expectation β. Let further $\bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i$ and $\bar{S} = \frac{1}{m}\sum_{i=1}^{m} S_i$.

It is of interest to combine the information from the two data bases to a common estimator for β. Two estimators are suggested:

$$\hat{\beta}_1 = \frac{1}{2}\bar{T} + \frac{1}{2}\bar{S} \qquad \text{and} \qquad \hat{\beta}_2 = \frac{\sum_{i=1}^{n} T_i + \sum_{i=1}^{m} S_i}{n+m}$$

a) Check for both estimators whether they are unbiased.
Calculate an expression for the variance of $\hat{\beta}_2$.
It can be shown that $\mathrm{Var}(\hat{\beta}_1) = (\beta^2/4)(1/n + 1/m)$. Which one is the best estimator when $n = 41$ and $m = 17$?

After closer considerations it is decided not to use the data from the second data base as these data are from detectors operating under working conditions quite different from those relevant for the current study. In the rest of the problem we shall thus only use the lifetimes $T_1, T_2, \ldots, T_n$ reported to the first data base.

b) Show that the likelihood function for β can be written $L(\beta) = \frac{1}{\beta^n} e^{-\sum_{i=1}^{n} t_i/\beta}$, and derive the maximum likelihood estimator (MLE) for β.
Calculate the estimate when $n = 41$ and $\sum_{i=1}^{41} t_i = 107$.
Find the 95% Wald confidence interval for β and calculate the interval numerically when the data are as given above.
Comment briefly why we can use a Wald confidence interval in this situation.

Information about β is also available from other sources than the data used above (e.g. expert knowledge, information from use under different conditions etc) - and we will now add this information using a Bayesian approach.

As prior distribution for β we use an inverse gamma distribution which has probability density function

$$p(\beta) = \frac{1}{b^a \Gamma(a)} \beta^{-a-1} e^{-1/(\beta b)} \quad , \quad \beta > 0$$

In this distribution the expectation is $1/(b(a-1))$ and variance is $1/(b^2(a-1)^2(a-2))$.

Moreover, for the quantiles of this distribution we have that $\xi_{1-\alpha/2,a,b} = \frac{2}{b\chi^2_{\alpha/2,2a}}$ for the $1-\alpha/2$ quantile (where $\chi^2_{\alpha/2,2a}$ is the $\alpha/2$ quantile in the χ^2_{2a}-distribution) and similar $\xi_{\alpha/2,a,b} = \frac{2}{b\chi^2_{1-\alpha/2,2a}}$ for the $\alpha/2$ quantile.

c) Find the posterior distribution and find an expression for the Bayes estimate for β. Calculate the estimate when $a = 4$, $b = 0.11$ and the data are as given in b).
Find a 95% Bayes interval for β. Compare this interval to the confidence interval in b) and comment briefly.