University of Stavanger

EXAM IN: STA500 INTRODUCTION TO PROBABILITY AND STATISTICS 2

DURATION: 4 HOURS DATE: FEBRUARY 10, 2014

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 2 PROBLEMS ON 4 PAGES, 9 PAGES OF ENCLOSURES.

Problem 1:

In this problem we shall consider a computer which has four central processing units (four "cores"). Each of these processing units can execute jobs independently of each others. We shall consider a situation where the execution time for each job is exponentially distributed with expectation 10 seconds, and the jobs arrive as a Poisson process with rate $\lambda = 20$ jobs per minute.

Further we assume (a bit simplified) that each new job which arrives is immediately sent to an idle processing unit if there are idle processors, and when all processing units are busy new jobs have to wait in queue.

a) What is the probability that exactly 25 jobs arrive during one minute?
What is the probability that more than 25 jobs arrive during one minute?
What is the probability that the execution time for a job is less than 5 seconds?

b) Is the queue stable? For which values of λ will the queue not be stable?
Consider a situation where there are 9 jobs in the system, i.e. 4 jobs are being executed and 5 jobs are waiting, when a new job arrives. What is the expected time until the new job starts being executed? What is the expected time until the new job is finished being executed?

Let $X(t)$ be the total number of jobs in the system (being executed or waiting in queue) at time t.

c) Explain briefly why $\{X(t):\ t \geq 0\}$ is a birth and death process.
Draw the transition graph for the process and write the transition rates on the plot. Include at least the seven first states (the states 0 to 6) in your plot.
Set up the steady state equations for the five first states (the states 0 to 4).
(You do not need to solve the equations.)

It can be shown that the steady state probabilities are $\pi_0 = 27/1267$, $\pi_1 = 90/1267$, $\pi_2 = 150/1267$ and $\pi_k = (500/3801) \cdot (5/6)^{k-3}$ for $k = 3, 4, 5, \ldots$.

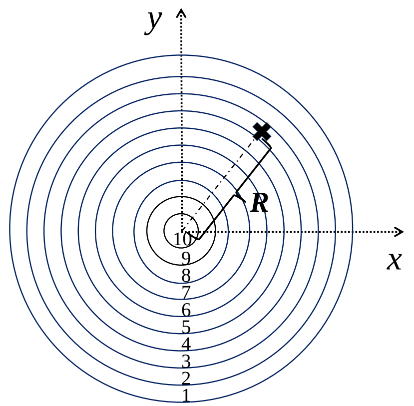
d) What is the long run proportion of time when all the four processing units are idle?
What is the long run proportion of time when all the four processing units are busy?
What is the expected number of jobs in the system in the long run? (Hint: The following might be useful: $\sum_{k=3}^{\infty} ka^k = a/(1-a)^2 - a - 2a^2$ when $a < 1$.)

Problem 2:

In this problem we shall consider the problem of hitting the center of a circular target. This applies to various situations like e.g. a helicopter aiming at hitting the center of the landing zone, a crane trying to place equipment at the correct spot and various sports like shooting, archery and darts. We shall use archery as our specific example. In this sport the practitioners shoot an arrow towards a circular target, and the closer to the center the arrow hits the more points are given. The target is divided into 10 circular zones, with 10 points given for hitting the inner zone and one point less for each zone away from the center. See the figure below.

Let R denote the distance from the point where the arrow hits to the center of the target. Then R can be expressed $R = \sqrt{X^2 + Y^2}$ where X and Y denotes respectively the horizontal and the vertical distance from the hitting point to the center (Pythagoras), see the figure. We assume that X and Y are independent and have a normal distribution with expectation 0 and the same standard deviation σ.

A sketch of an archery target with the ten zones, plus a coordinate system added on top in dotted lines, is given below.



a) Explain briefly, using results from the course/collection of formulas, that $R^2/\sigma^2 = X^2/\sigma^2 + Y^2/\sigma^2$ has a χ^2_2-distribution. (Hint: If $Z \sim N(0,1)$ then $Z^2 \sim \chi^2_1$)
Verify briefly by using a transformation result that R has a Weibull distribution with parameters $\alpha = 1/(2\sigma^2)$ and $\beta = 2$.
Show that the probability density function of R then becomes:

$$f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, \qquad r > 0$$

This special case of the Weibull distribution is called the Rayleigh distribution.

Kari is an eager archer, and she shoots on an archery target which has a radius of 40 cm (each zone is then 4 cm wide). Assume that it is known that for Kari is $\sigma = 14$. Also assume that the results of consecutive shots are independent.

b) Show that $F_R(r) = 1 - e^{-\frac{r^2}{2\sigma^2}}$.
Show that the probability that Kari hits the inner zone of radius 4 cm (i.e. gets 10 points) in a shot is 0.04.
Find the probability that Kari misses the target in a shot (i.e. the arrow goes outside the target).

Kari starts to shoot, and she continues until she hits the inner zone (i.e. gets 10 points) the first time. Let W be the number of shots she needs to shoot until she hits the inner zone the first time.

c) Which probability distribution does W have? Explain why.
Calculate $P(W < 5)$ and $\mathrm{E}(W)$.

Kari is going to shoot m arrows, and we shall in particular look at the worst of these m shots, $V = \max(R_1, \ldots, R_m)$.

d) Find an expression for the cumulative distribution function of V.
Use this to calculate the probability that the worst of $m = 25$ shots goes outside the target.
Explain briefly how you could alternatively have calculated this probability by using a binomial distribution (you do not need to do the calculation).

Karis sister Hanna has gotten interested in archery as well. After having trained for a while she claims that she is better than Kari. To prove this she will measure the distance from the center to the hitting point for n independent shots, $R_1, \ldots, R_n$, and use this informationen to estimate her σ.

In the rest of the problem we shall consider Hannas σ and assume this to be unknown.

e) Show that the maximum likelihood estimator (MLE) for σ becomes

$$\hat{\sigma} = \sqrt{\frac{1}{2n} \sum_{i=1}^{n} R_i^2}$$

Calculate the estimate when $n = 50$ measurements gave $\sum_{i=1}^{50} r_i = 785$ and $\sum_{i=1}^{50} r_i^2 = 16017$.

f) Find a 90% confidence interval for σ when the data are as given in point e). (Hint: This can be done in several ways. Pick a suitable approach, and explain why you can use the approach you have choosen.)
Does it look like Hanna is better than Kari?