Assignment

Statistics 302 Assignment 2 Report written by: Number: ” I confirm that I have not received help from, or given help to, anyone else in constructing the

solution to this assignment”

____

Name

Q1. (10) The number of UBC students infected by H1N1 is a Poisson random variable with rate 5 per week.

(a) What is the probability that there are more than 10 students infected in a week?

Solution:

The poisson probability is given by:

p(x; μ) = (e-μ) (μx) / x!

The probability that more than 10 students are infected in a week =

P(x> 10, μ = 5) = e-(5) * [(511)/11! + (512)/12! + (513)/13! + … + 5^20/20! +..]

P(x> 10, μ = 5) = e-(5) * 2. 0325

P(x> 10, μ = 5) = 0. 01369

(b) What is the probability that in the next 4 weeks there will be at least 2 weeks with more than 10 infected students each?

Solution:

The probability that in the next 4 weeks there will be at least 2 weeks with more than 10 students infected.

In a week the probability is 0. 01369, thus in the next 4 weeks there will be at least 2 weeks with more than 10 students infected is given by:

P = P_12 P_13 P_23 P_123 P_14 P_24 P_124 P_34 P_134 P_234 P_1234

P = 0. 0136911 = 3. 165468013 x 10-21

Week 1

Week 2

Week 3

Week 4

0

0

0

0

1

0

0

0

0

1

0

0

1

1

0

0

0

0

1

0

1

0

1

0

0

1

1

0

1

1

1

0

0

0

0

1

1

0

0

1

0

1

0

1

1

1

0

1

0

0

1

1

1

0

1

1

0

1

1

1

1

1

1

1

Q2. (10) (a) Suppose X is Binomial(n, p), with probability mass function f (k; n, p). show that

Solution:

a) Proof that f(k; n, p) / f(k − 1; n, p) = 1 + [(n + 1)p − k / k(1 − p)]:

nCk* Pk* (1-P)n–k / nCk-1* Pk-1* (1-P)n–k+1

= n! * (n-k)! *(n-k+1) * (k-1)! * Pk * P * (1-P)n-k / n! * (n-k)! * (k-1)! * k * P k * (1-P)n-k * (1-P)

= (n-k+1) * P / (1-P) * k

= nP-kP+P / k(1-P)

= k(1-P)+(n+1)P-k / k(1-P)

= 1+ [(n+1)P-k] / k(1-P)

(b) Suppose X is Binomial(n, p) and Y is Binomial(n, 1 − p), show that

Solution:

b) Proof that: P(X >= k) = P(Y = k)

=[nCk+nCk+1+nCK+2+…+nCn] * Pk+k+1+K+2+…+n * (1-P)n-k+n-(k+1)+n-(K+2)+…+1

Given the fact that nCk = nCn-k the first term of the multiplication can be written as:

[nC0+nC1+nC2+…+ nCn-k]

The second term in the multiplication can be written as:

Pn-0+n-1+n-2…+k

The Third term can be written as:

(1-P)1+2+3+…+n-k

Exchanging the places of the second and third terms, P(X >= k) can be written as:

[nC0+nC1+nC2+…+ nCn-k] * (1-P)1+2+3+…+n-k * Pn-0+n-1+n-2…+k

= P(Y 0.)

Solution:

The probability that three or more crashes will occur next year is given by:

P(x>= 3; μ = 1. 5)

= e-1. 5 * [1. 53/3! + 1. 54/4! + 1. 55/5! + ...+ 1. 513/13! +...]

= e-1. 5 * 0. 85668

= 0. 1911

Q5. (10) You are allowed to take a certain test three times, and your final score will be the maximum of the test scores. Thus,

X = max{X1 , X2 , X3},

where X1 , X2 , X3 are the three test scores and X is the final score. Assume that your test scores are (integer) values between 1 and 10 with equal probability 1/10, independently from each other. What is the probability mass function of the final score?

Solution:

We have 3 independent random variables X1 , X2 , X3 with uniform distribution, given by:

p1(x) = 1/10

p2(x) = 1/10

p3(x) = 1/10.

In order to calculate the probability mass function of the final score, we have:

= max{ p1, p2, p3 }

F(x) = P(max{ p1, p2, p3 } = x)

Q6. (10) The metro train arrives at the station, always on time, near your home every quarter hour starting at 6: 00 AM. You walk into the station every morning between 7: 10 and 7: 30 AM, with the time in this interval being a uniform random variable.

(a) What is the probability density function of the amount of the time, in minutes, that you have to wait for the first train to arrive?

Solution:

= 1/(0. 25 hours) = 1/15 per minute

U = (7. 1666, 7. 5), it takes values within this interval (hours)

F(U) = 1 / (7. 5*60 – 7. 166*60) = 0. 002272727272, uniform distribution in minutes.

The arrival time follows an exponential distribution, which is given by:

Arrive(x= 1 ; ✧ = 1/15) = F(U) ✧ e-✧ x

Arrive(x= 1 ; ✧ = 1/15) = 0. 002272727272 e-(1/15) (1/15) = 0. 0001417434825

The probability to wait is small.

(b) What is the expected waiting time?

Solution:

E(x) = F(U) (1/✧) = 0. 002272727272 15 = 0. 03409090909 minutes

(c) What is the median waiting time?

Solution:

Median(x) = F(U) ln(2)/ ✧ = 0. 002272727272 ln(2) 15 = 0. 02363001751 minutes

(d) Which of the expected waiting time and the median waiting time is a better summary? Why?

Solution:

The median waiting time is better representation of the phenomena.

Because…

The mean is the most commonly-used measure of central tendency in a distribution. The mean is valid only for interval data. Since it uses the values of all of the data points in the populations or sample, the mean is influenced by outliers that may be at the extremes of the data set. On the other hand, the median can be determined for ordinal data as well as interval data. Unlike the mean, the median is not influenced by outliers at the extremes of the data set.

Q7. (10) Suppose that the probability density function of a random variable X is as follows:

(a) Find the value of t such that P (X ≤ t) = 1/4.

Solution:

Solution for t is given by: t = 2, note that negative solution is neglected.

(b) Find the value of t such that P (X ≥ t) = 1/2.

Solution:

Solution for t is given by: t = 8^(1/2) = 2. 8242, note that negative solution is negleted.

(c) After the value of X has been observed, let Y be the integer closest to X, and Y is set equal to zero in case of tie. Find the probability mass function of the random variable Y.

Solution:

Y(X) = int(X) = i for i = 1, 2, 3, …n

Y = 0, if tie.

Thus the probability mass function is given by:

Where we have substituted integrals by summation.

Q8. (10) A system consists of 3 components arranged in series. The lifetime (in days) of each component follows approximately an exponential distribution with a mean lifetime of 100 days. The lifetimes of the components are independent.

(a) What is the probability that the first component lasts between 50 and 100 days?

Solution:

The exponential probability is given by:

p(x; ✦) = ✦ e-✦X

Where λ is the rate parameter. $\lambda = 1/100$

(b) What is the probability that 2 of the 3 components have lifetimes between 50 and 150 days?

Solution:

The exponential probability is given by:

$p(x; \lambda) = \lambda e^{-\lambda X}$ for one component.

$P(50 < x < 150 ; \lambda = 1/100)$, because are independent components.

(c) What is the cumulative distribution function of the lifetime of the entire system? What are the corresponding median and mean lifetime?

Solution:

The cumulative distribution function of the lifetime of the entire system is given by:

Thus, the cumulative function of the 3 components is given by.

where $\lambda = 1/100$,

(d) If the 3 components are arranged in parallel, what is the c. d. f. of the lifetime of the entire system? What is the corresponding median lifetime?

Solution:

The probability of the system in parallel is given by:

Where $p1(x, \lambda)$, $p2(x, \lambda)$, $p3(x, \lambda)$ are three independent exponential distributions in parallel. After some algebra and if we take q as any of the three independent exponential distributions, q = p1 or q = p2, q = q3, $q = \lambda e^{-\lambda X}$:

Thus, the cumulative distribution function of the lifetime of the entire system in parallel is given by:

It is the cumulative function of the entire system in parallel.

What is the corresponding median lifetime?

It is given by:

where ✧ = 1/100, the median is given solving numerically the above equation, what is given by:

Median = 18. 41662750 days (This is the median lifetime of the entire system in parallel configuration)

The numeric solution was calculated using Derive 6 software.

Refereces

Derive 6 software.