PRACTICE Final

STAT 410 FALL 2020

Wednesday, December 9, 2020

Name & UIN ....

Instructions:

1. You should finish this exam in 3 hours.

2. You can use any result shown in the lectures but you have to clearly state the result here.

3. Show all your supporting work. If you just write the answer you may not get a lot of credit.

4. There are 8 questions each worth 20 points. The maximum you can score is 160. This exam is worth 150 points and the remaining 10 points are for extra credit.

Problem 1: (6 + 6 + 8 points)

Let $N(t)$ be a Poisson process with rate λ. Let $X = N(t_1), Y = N(t_2) - N(t_1)$ and $Z = N(t_3) - N(t_2)$ where $0 < t_1 < t_2 < t_3$ are some fixed times.

a) What is the conditional distribution of X given $X + Y + Z = 10$.

b) What is the conditional joint distribution of X, Y, Z given $X + Y + Z = 10$? Is it a named distribution we have studied?

c) What is the conditional distribution of $Y + Z$ given $X + Y + Z = 10$?

(a) $$P(X=x \mid X+Y+Z=10) = \frac{P(N(t_1)=x)\, P(N(t_3)-N(t_1)=10-x)}{P(N(t_3)=10)}$$

$$= \frac{\frac{(\lambda t_1)^x}{x!} \cancel{e^{-\lambda t_1}} \frac{(\lambda(t_3-t_1))^{10-x}}{(10-x)!} \cancel{e^{-\lambda(t_3-t_1)}}}{\frac{(\lambda t_3)^{10}}{10!} \cancel{e^{-\lambda t_3}}}$$

$$= \frac{10!}{x!(10-x)!} \left(\frac{t_1}{t_3}\right)^x \left(1-\frac{t_1}{t_3}\right)^{10-x}$$

$$\Rightarrow X \mid X+Y+Z=10 \sim Bin\left(10, \frac{t_1}{t_3}\right)$$

(b) $$P(X=x, Y=y, Z=10-x-y \mid X+Y+Z=10)$$

$$= \frac{P(N_1(t_1)=x)\, P(N(t_2-t_1)=y)\, P(N(t_3-t_2)=10-x-y)}{P(N(t_3)=10)}$$

$$= \frac{10!}{x!\, y!\, (10-x-y)!} \left(\frac{t_1}{t_3}\right)^x \left(\frac{t_2-t_1}{t_3}\right)^y \left(\frac{t_3-t_2}{t_3}\right)^{10-x-y}$$

multinormal Distribution: $Mult\left(10, \left(\frac{t_1}{t_3}, \frac{t_2-t_1}{t_3}, \frac{t_3-t_2}{t_3}\right)\right)$.

(c) $$P(Y+Z=k \mid X+Y+Z=10) = P(X=10-k \mid X+Y+Z=10).$$

$$= \frac{10!}{(10-k)!\, k!} \left(\frac{t_1}{t_3}\right)^{10-k} \left(1-\frac{t_1}{t_3}\right)^k$$

$$\sim Bin^2\left(10, \frac{t_3-t_1}{t_3}\right).$$

Problem 2:(10 + 10 points)

Let W be double exponential random variable with pdf $f(x) = \exp(-|x|)/2$. A measuring device is used to observe W, but the device can only handle positive values, and gives a reading of 0 if $W < 0$; this is an example of censored data. So assume that $X = W\mathbb{I}(W > 0)$ is observed rather than W where $\mathbb{I}(W > 0)$ is the indicator of $W > 0$. Find $E(X)$ and $Var(X)$.

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & x = 0 \\ \frac{1}{2} + \int_0^x \frac{1}{2} e^{-t}\,dt = 1 - \frac{1}{2}e^{-x}, & x > 0 \end{cases}$$

$$f_X(x) = \begin{cases} 0, & x \le 0 \\ \frac{1}{2}e^{-x}, & x > 0 \end{cases}$$

$$E(X) = \int_0^\infty \frac{1}{2} e^{-x}\, x\, dx = -\int_0^\infty \frac{1}{2} x\, de^{-x}$$

$$= -\frac{1}{2} x e^{-x}\Big|_0^\infty + \int_0^\infty \frac{1}{2} e^{-x} dx.$$

$$= -\frac{1}{2} e^{-x}\Big|_0^\infty = \frac{1}{2}$$

$$Var(X) = E(X^2) - E^2(X)$$

$$= \int_0^\infty \frac{1}{2} e^{-x} x^2 dx - \frac{1}{4} = -\int_0^\infty \frac{1}{2} x^2 de^{-x} - \frac{1}{4}$$

$$= -\frac{1}{2} x^2 e^{-x}\Big|_0^\infty + \int_0^\infty x e^{-x} dx - \frac{1}{4}$$

$$= \frac{3}{4}$$

Problem 3:(6 + 6 + 8 points)

A drunken man wanders around randomly in a large space. At each step, he flips a fair coin and chooses to move either horizontally or vertically. Then he moves W_n steps horizontally or vertically (depending on the coin toss) where $W_1, W_2, \ldots, W_n, \ldots$ is a sequence of i.i.d discrete random variables supported on the non negative integers. Let (X_n, Y_n) and R_n be his position and distance from the origin after n steps, respectively.

(a) Determine whether or not X_n is independent of Y_n. Give reasoning.

(b) Find $Cov(X_n, Y_n)$ in terms of EW_1.

(c) Find $E(R_n^2)$ in terms of $E[(W_1)^2]$.

$X_n = \sum_{j=1}^{n} I_j W_j \quad Y_n = \sum_{j=1}^{n} (1-I_j) W_j$

(a) No. $X_n + Y_n = \sum_{i=1}^{n} W_i$ $E(X_n \mid Y_n = 0) = nEW_1$

$E(X_n \mid Y_n = 1) = nEW_1 - 1.$

(b) prob of one step be horizontal and be vertical are both 0.5. $\Rightarrow E(X_n) = E(Y_n) = \frac{1}{2} E\left(\sum_{i=1}^{n} W_i\right) = \frac{n}{2} EW_1$

$\Rightarrow Cov(X_n, Y_n) = Cov\left(\sum_{j=1}^{n} I_j W_j, \sum_{i=1}^{n} (1-I_i) W_i\right)$

$= \sum_{j=1}^{n} Cov\left(I_j W_j, \sum_{i=1}^{n} (1-I_i) W_i\right)$

$Cov(I_j W_j, (1-I_i) W_i) = 0$, if $i \neq j$.

$= \sum_{j=1}^{n} Cov(I_j W_j, (1-I_j) W_j)$

$= n\, E\left(I_j W_j - \frac{1}{2} EW_1\right)\left((1-I_j) W_j - \frac{1}{2} EW_1\right)$

$= n\left(E\left(I_j (1-I_j) W_j^2\right) - \frac{1}{4} (EW_1)^2\right)$

$= n\left(0 - \frac{1}{4} (EW_1)^2\right) = -\frac{n}{4} (EW_1)^2$

(c) $(X_n + Y_n)^2 = X_n^2 + Y_n^2 + 2X_nY_n = \left(\sum_{i=1}^{n} W_i\right)^2$

$E(X_n^2 + Y_n^2) = E\left(\sum_{i=1}^{n} W_i\right)^2 - 2E(X_nY_n)$

$= nE(W_1^2) + (n^2 - n)(EW_1)^2 - 2\left[Cov(X_n, Y_n) + E(X_n)E(Y_n)\right]$

$= nE(W_1^2) + (n^2 - n)(EW_1)^2 - 2\left[-\frac{n}{4} (EW_1)^2 + \frac{n^2}{4} (EW_1)^2\right]$

$$E(X_n^2+Y_n^2) = 2E(X_n^2) = 2E\left(\sum_{j=1}^{n} I_j W_j\right)^2 = 2\left(nEI_j^2 EW_j^2 + \frac{n^2-n}{4}(EW_1)^2\right)$$

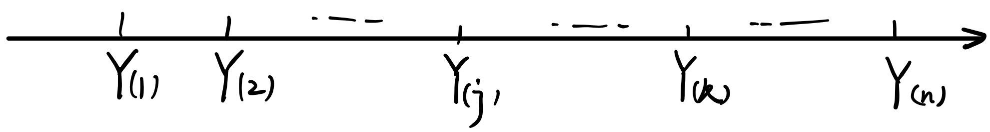
Problem 4: (10 + 10 points)

We have observed random variables $Y_1, Y_2, \ldots, Y_n$ i.i.d $N(0,1)$ random variables.

(a) For fixed j and k with $1 \le j < k \le n$ find $P(Y_{new} \in [Y_{(j)}, Y_{(k)}])$.

(b) Does your answer change if the Y_i are i.i.d $Unif(0,1)$ random variables? Does your answer change if the Y_i are i.i.d $Bernoulli(0.5)$ random variables?

(a)



$$P(Y_{new} \in [Y_{(j)}, Y_{(k)}] = \frac{k-j}{n+1}$$

(b) Doesn't chage when $Y_i \sim Unif(0,1)$.

$$P(Y_{new} \in [Y_{(j)}, Y_{(k)}] \equiv 1.$$

Chage when $Y_i \sim Bern(0.5)$.

Problem 5: (5 + 5 + 5 + 5 points)

Let $X_1, X_2, \ldots, X_n$ and Y be i.i.d continuous random variables with CDF F. Let $V_i = \mathbb{I}(X_i < Y)$.

(a) Are V_i independent? Are V_i conditionally independent given Y? Explain.

(b) Let $N = \sum_{i=1}^{n} V_i$. Find EN.

(c) Find the conditional distribution of N, given $Y = y$

(d) Find $Var(N)$.

(a) No V_i depend on Y. , Yes

$P(V_1 = 1 \mid V_2 = 1) > P(V_1 = 1 \mid V_2 = 0)$.

(b) $EN = n EV_i = \frac{n}{2}$

(c) $V_i \mid Y = y = I(X_i < y) \sim Bern(F(y))$.

$\Rightarrow N \mid Y = y \sim Bin(n, F(y))$.

(d) $Var(N) = E(Var(N \mid Y)) + Var(E(N \mid Y))$.

$= E(nF(y)(1 - F(y))) + Var(nF(y))$.

$= \frac{n}{2} - \frac{n}{3} + \frac{n^2}{12} = \frac{n^2}{12} + \frac{n}{6}$

Problem 6: (6 + 6 + 8 points)

Let Z_1, Z_2 be $N(0,1)$ i.i.d random variables.

(a) As a function of Z_1, Z_2, create an $Gamma(2,1)$ r.v. X (your answer can also involve the standard Normal CDF).

(b) Let $Y = \exp(-R)$, where $R = \sqrt{Z_1^2 + Z_2^2}$. Write down (but do not evaluate) an integral for $E(Y)$.

(c) Let $X_1 = Z_1 + Z_2$ and $X_2 = Z_1 - Z_2$. Determine whether X_1 and X_2 are independent.

(a) C.D.F of exp(1) $F(x) = 1 - e^{-x} \quad x > 0.$

$\Rightarrow F^{-1}(y) = -\log(1-y),\ y \in (0,1)$

$F(X) \sim Unif(0,1)$

$F^{-1}(\Phi(Z_i)) \sim exp(1). \quad i = 1,2$

$\Phi(Z_i) \sim Unif(0,1) \Rightarrow 1 - \Phi(Z_i) \sim Unif(0,1)$

$\Rightarrow -\log(\Phi(Z_i)) \sim exp(1)$

$\Rightarrow -\log(\Phi(Z_1)) - \log(\Phi(Z_2)) \sim Gamma(2,1)$

(b) $E(Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\sqrt{z_1^2 + z_2^2}} \cdot \frac{1}{2\pi} e^{-\left(\frac{z_1^2}{2} + \frac{z_2^2}{2}\right)} \, dz_1 dz_2$

(c)

Problem 7: (6 + 6 + 8 points)

A Markov chain $X_0, X_1, \ldots$ with state space $\{-2, -1, 0, 1, 2\}$ proceeds as follows. The chain starts at $X_0 = 0$. If X_n is not an endpoint 2 or -2, then X_{n+1} is $X_n + 1$ or $X_n - 1$, each with probability 1/2. Otherwise, the chain gets reflected of the endpoint, i.e., from 2 it always goes to 1 and from -2 it always goes to -1.

(a) Is $|X_0|, |X_1|, |X_2|, \ldots$ also a Markov chain? Explain.

(b) (b) Let sgn be the sign function: $sgn(x) = 1$ if $x > 0$, $sgn(x) = -1$ if $x < 0$, and $sgn(0) = 0$. Is $sgn(X_0), sgn(X_1), sgn(X_2), \ldots$ a Markov chain? Explain.

(b) Find the stationary distribution of the chain $X_0, X_1, X_2, \ldots$.

(a) Yes $|X_i|$ only depends on $|X_{i-1}|$.

(b). no $P(Sgn(X_3)=0 \mid sgn(X_2)=-1,$
$sgn(X_2)=-1$
$sgn(X_0)=0)=0.$

$P(Sgn(X_3)=0 \mid sgn(X_2)=-1)>0.$

(C) -2 -1 0 1 2.

degrees: 1 2 2 2 1.

stationary distribution $\propto$ degrees of nodes.

$$\Rightarrow S = \left(\frac{1}{8}, \frac{2}{8}, \frac{2}{8}, \frac{2}{8}, \frac{1}{8}\right).$$

Problem 8:(10 + 10 points)

(a) Let $X \sim Exp(\lambda)$ and $Y|X = x \sim Poi(x)$. What is the distribution of Y?

Let N_t be a Poisson process with rate λ. Define τ to be the time of first arrival in the process N_t, so that $\tau = \min\{t : N_t = 1\}$.

(b) What is the distribution of $N_{2\tau} - N_\tau$?

(a). Consider two Possion process I : rate 1
II : rate λ
begin at a same time.

Let X be the time of first event belong to II.
Y be the number of events belog to I before the time of the first event belongs to II.

$Y \sim Geom(\frac{\lambda}{\lambda+1})$.

(b) $N_{2T} - N_T \,|\, T \sim Pois(T)$.

$T \sim Expo(\lambda)$

$\Rightarrow N_{2T} - N_T \sim Geom(\frac{\lambda}{\lambda+1})$.