Stat310: Test 2

Answer the questions in the spaces provided on the question sheets. If you run out of room for an answer, continue on the back of the page. You have **two hours** and you may use one double-sided page of notes, and wolframalpha if you have to. You may not use your text book, or communicate anything about the contents of this test to anyone else. Please pledge on the last page.

Name: ______________________________

1. Let $X_1 \sim \text{Exp}(\theta)$, $X_2 \sim \text{Exp}(\theta)$ and X_1 and X_2 are independent. (Hint: $\Gamma(2) = 1$)

   (a) Using bivariate change of variables, find the pdf of $S_2 = X_1 + X_2$. What named distribution does this pdf represent? (7)

$A = X + Y \qquad B = X$ 1 $\qquad f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{\theta^2} e^{-\frac{(x+y)}{\theta}}$ 1

$X = B \qquad Y = A - B$ 1

$x > 0 \Rightarrow b > 0$

$y > 0 \Rightarrow a - b > 0 \Rightarrow b > a$ 1

$$J = \begin{vmatrix} \frac{\partial x}{\partial a} & \frac{\partial x}{\partial b} \\ \frac{\partial y}{\partial a} & \frac{\partial y}{\partial b} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$ 1

$$f_{A,B}(a,b) = f_{X,Y}(a-b,\ b)\ |J| = \frac{1}{\theta^2} e^{-\left(\frac{a-b+b}{\theta}\right)} = \frac{1}{\theta^2} e^{-a/\theta}$$ 1

$$f_A(a) = \int_0^{a} \frac{1}{\theta^2} e^{-a/\theta}\, db = \frac{a}{\theta^2} e^{-a/\theta} = \frac{\theta^{-2}}{\Gamma(2)} a^{2-1} e^{-a/\theta}$$ 1

$\Rightarrow A \sim \text{Gamma}(\alpha = 2, \theta = \theta)$

this integral is harder than I said I'd give so no penalty for failing to integrate it

+1 bonus if you did get this far

(b) Let $X_3 \sim \text{Exp}(\theta)$ and be mutually independent with X_1 and X_2. What is the distribution of $S_3 = X_1 + X_2 + X_3$? Use mgfs to find out. (3)

$$M_{S_3}(t) = \left[M_{X_1}(t)\right]^3 \quad 1$$

$$= \frac{1}{(1-\theta t)^3} \quad 1$$

$$\Rightarrow \quad S_3 \sim \text{Gamma}(\alpha = 3, \theta) \quad 1$$

2. Let $H_i \sim \text{Normal}(\mu = 80, \sigma^2 = 25)$ be an iid sequence that represents your homework scores (as a percentage of 100).

(a) Give three reasons why your homeworks are not independent. For each reason, characterise its effect on the correlation between scores. (3)

* carry over effects - bad score < more effort next time / depressed and give up

* represent your underlying math/stat skills - you wouldn't expect them to be totally unrelated

* material gets harder over the semester => scores decrease

any 3 ✓

* material is cumulative

(b) Is the normal distribution an appropriate approximation to your homework scores? Why/why not? (2)

2pt - One reason why it is a poor fit. (e.g. range from $-\infty, \infty$, normal dist includes all real numbers, but homework scores stick to integers, negative values, not independent, etc.)

And one reason why it's still okay to use as an approximation anyways.

ex. the variance is small enough to keep the observed range reasonable.

the homework score is like a mean, which will be normal by CLT

Scores will be symmetrically clustered around a mean

1pt - A conclusion with one or more arguments supporting just that conclusion

(c) Ignoring the qualms above, let $T_n = \sum_{i=1}^{n} H_i$ be your total homework score after n homeworks. What is the distribution of T_n? Justify mathematically. (3)

$H_i \sim \text{normal}(80, 25)$ which is given

$$M_{T_n}(t) = E(e^{T_n t}) = E(e^{\sum_{i=1}^{n} H_i t}) = E\left(\prod_{i=1}^{n} e^{H_i t}\right) = \prod_{i=1}^{n} E(e^{H_i t})$$

$$= E(e^{H_i t})^n \quad \text{because } H_i \text{ are i.i.d.}$$

$$= M_{H_i}(t)^n$$

$$= \left(e^{\mu t + \frac{1}{2}\sigma^2 t^2}\right)^n$$

$$= e^{n\mu t + \frac{1}{2}\sigma^2 t^2}$$

+1 pt proof

let $\mu' = n\mu$ and $\sigma' = \sqrt{n}\,\sigma$ then we can recognize this as the mgf of a $\text{Normal}(n\mu, n\sigma^2)$ random variable.

$$\therefore T_n \sim \text{Normal}(80n, 25n)$$

+1 pt normal

+1 pt parameters

(d) Let $\bar{H}_n = T_n/n$, your average homework score. What assumptions can you relax so that $\bar{H}_n$ is approximately normally distributed? Why? (2)

We can relax our assumption that the H_i are distributed normally, because the CLT guarantees that the sample mean $\bar{H}_n$ will have an assymptotically normal distribution no matter what the underlying distribution.

normality = 1pt
CLT = 1pt

3. Let $f(x,y) = c(x^2+y^2)$, $x \in [0,1]$, $y \in [0,1]$. This pdf might represent your scores on the first and second homework as proportions between 0 and 1, no longer assuming independence.

(a) What value of c makes f a valid bivariate pdf? What other properties do you need to check? (4)

Need to check

2pts

1. $\underline{\int_0^1\int_0^1 f(x,y)\,dx = 1}$ and $\underline{f(x,y) \geq 0 \ \forall x \in [0,1] \text{ and } y \in [0,1]}$

2. $\int_0^1\int_0^1 c(x^2+y^2)\,dx\,dy = 1$ ← +1 pt

$2\int_0^1 cx^2\,dx = 1$ by symmetry

$2c\,\frac{x^3}{3}\Big|_0^1 = 1$

$\frac{2}{3}c = 1$

$\therefore c = \frac{3}{2}$ ← +1pt

3. $\frac{3}{2}(x^2+y^2) \geq 0 \ \forall x \in [0,1] \ \& \ y \in [0,1]$ because the function is even in x and y

(b) What is $P(X > 0.5 \cap Y > 0.5)$, the probability you get over 50% on both homeworks? (3)

$$= \frac{3}{2} \int_{0.5}^{1}\int_{0.5}^{1} x^2 + y^2 \; dy \; dx$$

1pt 1pt

$$\approx 0.4375$$ 1pt

(c) What is $E(\frac{X+Y}{2})$, the average grade on homeworks one and two? (3)

$$= \frac{1}{2} E[X+Y]$$

$$= \frac{1}{2} \cdot \frac{3}{2} \int_{0}^{1}\int_{0}^{1} (x+y)(x^2+y^2) \; dy \; dx$$

1pt 1pt

$$= 5/8 \approx 62.5\%$$ 1pt

(d) Outline a strategy that would enable you to find $P(\frac{X+Y}{2} > 0.5)$. (2)

* transformation + univariate integral 2pt

OR

* trickier bivariate integral 2pt

4. (a) Let $B_n \sim \text{Binomial}(n = n, p = \lambda/n)$, i.e. a binomial distribution where the number of successes is constant, but the size grows. Complete the blanks to find $\lim_{n\to\infty} B_n$. A helpful mathematical fact is that $\lim_{n\to\infty}(1 + \frac{x}{n})^n = e^x$ (8)

$$P(B_n = x) = \underline{\binom{n}{x}\left(\frac{x}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}} \quad \text{1pt}$$

pmf of binomial

$$= \frac{n!}{(n-x)!\,x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$

def. of $\binom{n}{x}$ ½

$$= \underbrace{\frac{n!}{n^x(n-\cancel{k})!}}_{①}\underbrace{\left(\frac{\lambda^k}{k!}\right)}_{②}\underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{③}\underbrace{\left(1-\frac{\lambda}{n}\right)^{-x}}_{④}$$

$$\lim_{n\to\infty} ① = \lim_{n\to\infty}\frac{n(n-1)(n-2)\cdots(n-\cancel{k}+1)}{n^{\cancel{k}}} = \lim_{n\to\infty}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\cdots\left(\frac{n-k+1}{n}\right) \quad \text{1pt}$$

$$= \underline{\ 1\ }$$

$$\lim_{n\to\infty} ② = \underline{\lambda^k/k!} \quad ½$$

$$\lim_{n\to\infty} ③ = \underline{e^{-\lambda}}$$ helpful mathematical fact 1pt

$$\lim_{n\to\infty} ④ = \underline{\ 1\ } \quad ½$$

$$\Rightarrow \lim_{n\to\infty} P(B_n = x) = \underline{\frac{\lambda^x}{x!}e^{-\lambda}} \quad \text{1pt}$$

$$\Rightarrow \lim_{n\to\infty} B \sim \text{Poisson}(\underline{\ \lambda\ }) \quad ½$$

Pledge (including time started and finished):

Question	Points	Score
1	10	10
2	10	10
3	12	12
4	8	8
Total:	40	40

:)