MATH 4710 **FINAL (SOLUTION)** **SPRING 2019**

Name :________________________________ Cornell NetID:________________________

Instructions

(1) Write clearly using a black or blue pen or pencil. Provide reasons for all your answers and explain your computations. For numerical answers, give either a simplified fraction or a decimal answer, which ever comes more easily.

(2) There are six completely independent problems and you have 150 minutes (2 hours and a half). The problems are arranged from the simplest to the more sophisticated. There are two blank pages at the end if you need more space.

(3) Keep your cell phone away. Do not use electronic devices. No books or notes are permitted. No communications with anyone during the final.

(4) Academic integrity is expected of all Cornell University students at all times, whether in the presence or absence of members of the faculty.

Problem 1 (7pts): An urn contains 5 balls, three white balls and two red balls. We pick two balls without replacement and put them in a bag. From that bag, we sample a ball, twice, with replacement.

(a) (2 pts) What is the probability that the two balls picked in the urn have the same color?

(b) (2 pts) What is the probability that the two samples taken from the bag have the same color?

(c) (3 pts) Given that the two samples taken from the bag have the same color, what is the probability that the two balls in the bag have the same color?

(a) (Imagine all 5 balls in the urn are numbered) The number of ways two pick two balls without replacement is $\binom{5}{2} = 10$. The number of pairs made of the same color is $\binom{3}{2} + 1 = 4$. The probability that the two balls picked in the urn have the same color is $\boxed{2/5}$.

(b) To compute this probability, we condition on whether or not the balls picked in the urn have the same color. The balls picked in the urn have the same color with probability $2/5$ and, in that case, the two balls picked in the bag are certain to be of the same color. The balls picked in the urn have probability $3/5$ to be of different colors and, in that case, the probability that the balls ticked in the bag have the same color is $1/2$.

Thus the probability that the two balls picked from the bag have the same color is $(2/5)\times 1+(3/5)\times(1/2)$ which equals $\boxed{7/10}$.

(c) Let B the event that the two balls in the bag have the same color and A the event that the balls sampled from the bag have the same color. We want to compute $P(B|A) = P(A\cap B)/P(A) = P(A|B)P(B)/P(A) = 1 \times (2/5)/(7/10)$. That is $\boxed{P(B|A) = 4/7}$.

Problem 2: (6 pts)

[Answers for this problem should be expressed using a small number of rationals n/m, the number e, and a small number basic operations $\times, \pm, /$. No numerical values because calculators are not allowed. Give short explanations for how you arrive at your answers.]

A certain typing agency employs 2 typists. The first makes an average of 3 errors per article. The second typist makes an average of 5 errors per article.

(a) (2 pts) Your next article is being typed by the first typist. What is your estimate that it will have at least one error?

(b) (2 pts) If a certain article is equally likely to typed by either typist, what is the estimated probability that it will contains no errors?

(c) (2 pts) The first typist is typing a certain typical article containing 1000 characters (including all symbols and spaces). We assume that the events that each character is typed correctly are mutually independent. Under this assumption, what is the probability distribution of the number Y of errors the typist will make while typing this article?

(a) We assume that the number of errors X is well approximated by a Poisson distribution with mean 3 which means the parameter $\lambda = 3$.

The probability of at least one errors is $\boxed{P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-3}}$.

(b) We proceed as before and condition on which typist types the article. The estimated probability of no error is $\boxed{\frac{1}{2}(e^{-3} + e^{-5})}$.

(c) The assumptions made indicate that the number Y of errors made by the first typist when typing this article is a Binomial random variable with parameter p and $n = 1000$. Where p is the probability that a given character is miss-typed. We now that $np = 3$ so $p = 3/1000$ so that $\boxed{Y \sim \mathcal{B}(3/1000, 1000)}$.

Problem 3: (6 pts) The random variables X, Y are independent standard normal $\mathcal{N}(0,1)$ random variables (recall the notation $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$). Set $U = 1 + 3X - Y$ and $Z = \min\{X, Y\}$.

(a) (2 pts) What is the variance of U?

(b) (2 pts) Find $P(Z < Y)$ and $P(Z > 0)$.

(c) (2 pts) Compute the density function f_Z of Z.

(a) We have $\mathrm{Var}(1 + 3X - Y) = \mathrm{Var}(3X - Y) = 9\mathrm{Var}(X) + \mathrm{Var}(Y)$ because X and Y are independent. Moreover $\mathrm{Var}(X) = \mathrm{Var}(Y) = 1$. So $\boxed{\mathrm{Var}(U) = 10}$.

(b) Note that $\{Z < Y\} = \{X < Y\}$. and $P(X < Y) + P(Y < X) = 1$ because $P(X = Y) = 0$. Also, $P(X < Y) = P(Y < X)$ because X, Y are independent with the same distribution. It follows that $\boxed{P(Z < Y) = 1/2}$.

Observe that $P(Z > 0) = P(X > 0 \text{ and } Y > 0) = P(X > 0)^2$ because X and Y are independent. Obviously, $P(X > 0) = 1/2$, so we get $\boxed{P(Z > 0) = 1/4}$.

We compute $P(Z \leq z) = 1 - P(Z > z) = 1 - P(X > z \text{ and } Y > z) = 1 - (1 - \Phi(z))^2$.

Then we take the derivative to obtain $\boxed{f_Z(z) = 2(1 - \Phi(z))\Phi'(z) = \frac{2}{\sqrt{2\pi}}(1 - \Phi(z))e^{-z^2/2}}$.

Problem 4: (7pts) The pair (X, Y) takes value in $\{0, 1, 2, \ldots\} \times \{0, 1, 2, \ldots\}$ and its joint mass distribution function is

$$p_{(X,Y)}(x, y) = \frac{e^{-\lambda}\lambda^y e^{-y} y^x}{x!y!}.$$

(a) (2 pts) What is the mass distribution p_Y of Y? Are X and Y independent or not?

(a) (2 pts) Find $P(X = 0)$.

(c) (3 pts) Find the conditional mass function of X given $Y = y$, $p_{X|Y}(x|y)$, and compute $E(X|Y = y)$. What is the random variable $E(X|Y)$? Compute $E(X)$.

(a) We compute $p_Y(y) = \frac{e^{-\lambda}\lambda^y e^{-y}}{y!} \sum_{x=0}^{\infty} \frac{y^x}{x!} = \frac{e^{-\lambda}\lambda^y e^{-y}}{y!} e^y$..
This gives $\boxed{p_Y(y) = e^{-\lambda}\frac{\lambda^y}{y!}}$ (Poisson λ). We compute

$$\frac{p_{(X,Y)}(x, y)}{p_Y(y)} = \frac{e^{-y}y^x}{x!}.$$

Because this is not a function of x alone, $\boxed{\text{the random variables } X, Y \text{ are not independent}}$.

(b) We have $\boxed{P(X = 0) = e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y e^{-y}}{y!} = e^{-\lambda} \sum_{y=0}^{\infty} \frac{(\lambda/e)^y}{y!} = e^{-\lambda+\lambda/e}}$.

(c) We have $\boxed{p_{X|Y}(x|y) = \frac{p_{X,Y)}(x,y)}{p_Y(y)} = \frac{e^{-y}y^x}{x!}}$ and $E(X|Y = y) = e^{-y} \sum_{x=0}^{\infty} x\frac{y^x}{x!}$. This is the expectation of a Poisson random variable with parameter y and it follows that $\boxed{E(X|Y = y) = y, E(X|Y) = Y}$. It follows that $\boxed{E(X) = E(E|Y)) = E(Y) = \lambda}$ because Y is a Poisson with parameter λ.

Problem 5: (8 pts)

Let X and Y be two independent exponential λ random variables. Please, treat each question (a), (b) and (c), independently of each other.

(a) (3 pts) Show that the random variable $U = \frac{X}{X+Y}$ is uniform on a certain interval.

(b) (3 pts) Find the density function of the pair (U, V) where $U = \frac{X}{X+Y}$ and $V = X + Y$. Are U and V independent?

(c) (2 pts) What is the probability that X is larger than $2Y$?

(a) We first observe that $U \in (0, 1)$. For $u \in (0, 1)$, we compute

$$P(U \leq u) = P(X \leq u(X+Y)) = P((X \leq \frac{u}{1-u}Y) = \lambda^2 \int_0^\infty \int_0^{uy/(1-u)} e^{-\lambda x} e^{-\lambda y} dx dy.$$

We have $\lambda \int_0^{uy/(1-u)} e^{-\lambda x} dx = 1 - e^{-\lambda uy/(1-u)}$ and

$$P(U \leq u) = \lambda \int_0^\infty e^{-\lambda y} - e^{-\lambda y - \lambda uy/(1-u)} dy = 1 - \lambda \int_0^\infty e^{-\lambda y/(1-u)} dy = u.$$

Recall that $F_U(u) = 0$ for $u \leq 0$ and $F_U(u) = 1$ for $u \geq 1$. This gives $\boxed{f_U(u) = \mathbf{1}_{(0,1)}(u)}$. The random variable U is uniform on $(0, 1)$.

(b) We use the change of variable formula to find the density function $f_{(U,V)}(u, v)$ using

$$\begin{cases} u = & x/(x+y) \\ v = & x+y \end{cases} \quad \text{and} \quad \begin{cases} x = & uv \\ y = & v(1-u) \end{cases}$$

between the regions $(x, y) \in (0, \infty) \times (0, \infty)$ and $(u, v) \in (0, 1) \times (0, \infty)$. This gives

$$J(u, v) = \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v(1-u) + uv = v.$$

It follows that, fro $u \in (0, 1)$ and $v \in (0, \infty)$, $f_{(U,V)}(u, v) = \lambda^2 e^{-\lambda(uv + v(1-u))} v = \lambda^2 v e^{-\lambda v}$.

This gives $\boxed{f_{(U,V)}(u, v) = \mathbf{1}_{(0,1)}(u) \times \lambda^2 v e^{-\lambda v} \mathbf{1}_{(0,\infty)}(v).}$ This is the product of a uniform density on $(0, 1)$ (for U) and a Gamma distribution (for V). It follows that $\boxed{U \text{ and } V \text{ are independent}}$.

(c) $P(X > 2Y) = \int_0^\infty \lambda e^{-\lambda y} \left\{ \int_{2y}^\infty \lambda e^{-\lambda x} dx \right\} dy = \int_0^\infty \lambda e^{-3\lambda y} dy.$

Because $\int_0^\infty 3\lambda e^{-3\lambda y} dy = 1$, $\boxed{P(X > 2Y) = 1/3}$.

Problem 6: (6 pts) Let $X_1, X_2, \ldots$ be a sequence of random variables (all defined on the same probability space). We do not assume they are independent. We set $S_n = X_1 + \cdots + X_n$.

We assume that $E(X_i) = i$, $\mathrm{Var}(X_i) = i$ and $\mathrm{Cov}(X_i, X_j) = \begin{cases} \min\{i,j\} & \text{if } |i-j| = 1 \\ 0 & \text{if } |i-j| > 1. \end{cases}$.

(a) (3 pts) Compute $\mu_n = E(S_n)$ and $\sigma_n^2 = \mathrm{Var}(S_n)$.

(b) (3 pts) Show that, for any $\epsilon > 0$, $\lim_{n\to\infty} P(|\frac{1}{\mu_n}(S_n - \mu_n)| \geq \epsilon) = 0$.

(a) $\mu_n = \sum_1^n E(X_i) = \sum_1^n i$. This gives $\boxed{\mu_n = n(n+1)/2}$.
Also, $\sigma_n^2 = \sum_1^n \mathrm{Var}(X_i) + 2\sum_{1\leq i<j\leq n} \mathrm{Cov}(X_i, X_j) = \sum_1^n i + 2\sum_1^{n-1} i = \frac{1}{2}(n(n+1) + 2(n-1)n)$.
This give $\boxed{\sigma_n^2 = n(3n-1)/2}$.

We compute $P(|\frac{1}{\mu_n}(S_n - \mu_n)| \geq \epsilon) = P(|S_n - \mu_n| \geq \epsilon\mu_n)$ and use Chebyshev's inequality

$$P(|S_n - \mu_n| \geq \epsilon\mu_n) \leq \frac{\sigma_n^2}{\epsilon^2\mu_n^2} = \frac{n(3n-1)}{2\epsilon^2(n(n+1))^2/4} = \frac{2(3n-1)}{\epsilon^2 n(n+1)^2}.$$

It follows that $\boxed{\lim_{n\to\infty} P(|\frac{1}{\mu_n}(S_n - \mu_n)| \geq \epsilon) = 0}$.