Statistics 630 – Exam I
Wednesday, 30 September 2020

Printed Name: ________________________ **Email:** ________________________

INSTRUCTIONS FOR THE STUDENT:

1. You have 50 minutes to complete the exam (after taking a moment to read these instructions). Please indicate your start time: ______________________ ,

   and your end time: ______________________
2. There are 6 pages including this cover sheet and the formula sheets.
3. Questions 1–4 are multiple choice and worth 5 points each.
4. Questions 5–8 require solutions to be worked out and are 10 points each. Please write out your answers in *the spaces provided*, explaining your steps. You may refer to theorems by name/description rather than by its number in the book.
5. If you *cannot* print out the exam, please write your answers on blank sheet of paper – in order.
6. You may use the *attached formula sheets*. No other resources are allowed. Do not use the textbook, the class notes, homework or formula sheets that were posted online.
7. *Do not use a calculator.* You may leave answers in forms that can easily be put into a calculator such as $\frac{12}{19}$, $\binom{40}{5}$, e^{-3}, $\Phi(1.5)$, etc.
8. Do not discuss or provide any information to anyone concerning any of the questions on this exam until your solutions are returned or I post my solutions.

I attest that I spent no more than 50 minutes to complete the exam. I used only the materials allowed above. I did not receive assistance from or provide assistance to anyone either before or while taking this exam.

Student's Signature ______________________________________

Questions 1–4 are multiple choice: circle the single correct answer. No partial credit!

1. (5 points) Consider the function $F(t) = \begin{cases} 0 & \text{if } t \le 0, \\ t^2 & \text{if } 0 < t < \frac{1}{2}, \\ 1-t & \text{if } \frac{1}{2} \le t < 1, \\ 1 & \text{if } t \ge 1. \end{cases}$

   $F(t)$ is <u>not</u> a cdf for some random variable X because

   (a) it is not continuous.
   (b) it is not nondecreasing.
   (c) it does not integrate to 1.
   (d) it is not a function of x.
   (e) all of the above.

2. (5 points) Suppose $W \sim \text{Poisson}(\lambda)$. We know that $\mathsf{E}(W) = \lambda$ and $\mathsf{Var}(W) = \lambda$. What is $\mathsf{E}(W(W-1))$?

   (a) $\lambda(\lambda - 1)$.
   (b) λ.
   (c) λ^2.
   (d) 2λ.
   (e) 0.

3. (5 points) A random vector (X, Y) has joint pdf $f(x,y) = \begin{cases} 5x^2y & \text{if } 0 \le y \le x \le 1, \\ 5xy^2 & \text{if } 0 \le x \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$

   Without any computation, we can say that

   (a) X and Y are independent because $f(x,y) = g(x)h(y)$ for some functions $g(x)$, $h(y)$.
   (b) X and Y are not independent because $f(x,y) = g(x)h(y)$ for some functions $g(x)$, $h(y)$.
   (c) X and Y are independent because $f(x,y) \ne g(x)h(y)$ for some functions $g(x)$, $h(y)$.
   (d) X and Y are not independent because $f(x,y) \ne g(x)h(y)$ for some functions $g(x)$, $h(y)$.
   (e) none of the above.

4. (5 points) Randy has 17 blue socks in his drawer, 11 of which have no holes, and 13 red socks, 5 of which have no holes. Randy pulls out two socks at random and sees they are both blue. What is the chance neither has a hole?

   (a) $\frac{11 \times 5}{17 \times 13}$.
   (b) $\frac{11 \times 10}{17 \times 16}$.
   (c) $\frac{\binom{11}{2}}{\binom{30}{2}}$.
   (d) $1 - \frac{\binom{6}{2}}{\binom{17}{2}}$.
   (e) $1 - \frac{\binom{19}{2}}{\binom{30}{2}}$.

Provide solutions to Questions 5–8, to the point of a calculable expression.

5. (10 points) The number of connections between a computer server and customer A is a Poisson(1.3) random variable and, independently, the number of connections with customer B is a Poisson(0.7) random variable. What is the chance the server has at least two connections with <u>each</u> of the two customers? (Show a computable expression.)

6. (10 points) At a given point in time, 2% of individuals in a community have COVID-19. Of those individuals, 0.8% are hospitalized in ICU with the disease. Of individuals without COVID-19, 0.01% are in ICU. If a patient in ICU is randomly selected, what is the chance they have COVID-19? (Show a computable expression.)

7. (10 points) Suppose Z has beta(2,2) pdf $f_Z(z) = 6z(1 - z)$ for $0 \leq z \leq 1$. What *equation must be solved* to get the .25-quantile (25^{th} percentile) for Z? Simplify the equation as much as possible, but you do not need to solve it.

8. (10 points) $Z \sim \text{gamma}(3, \lambda)$ and the conditional pdf for Y, given $Z = z$, is $f_{Y|Z}(y|z) = \frac{2y}{z^2}$, $0 \leq y \leq z$. Recall that the joint pdf for (Y, Z) is $f(y, z) = f_Z(z) f_{Y|Z}(y|z)$. Use this to find the marginal pdf for Y.

Formulas for Exam I

permutations $P_{n,k} = \frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1)$.

combinations $C_{n,k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$.

complement and union $\mathsf{P}(A^c) = 1 - \mathsf{P}(A)$; $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$.

conditional probability $\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}$.

independent events $\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$.

total probability $\mathsf{P}(A) = \sum_{k=1}^n \mathsf{P}(A \mid B_k)\mathsf{P}(B_k)$ if $B_1, \dots, B_n$ are disjoint, $\bigcup_{k=1}^n B_k = \Omega$.

Bayes' rule $\mathsf{P}(B_j \mid A) = \frac{\mathsf{P}(A|B_j)\mathsf{P}(B_j)}{\sum_{k=1}^n \mathsf{P}(A|B_k)P(B_k)}$ if $B_1, \dots, B_n$ are disjoint and $\bigcup_{k=1}^n B_k = \Omega$.

cdf of random variable $F_X(x) = \mathsf{P}(X \le x) = \sum_{y \le x} p_X(y)$ if X is discrete;
$F_X(x) = \mathsf{P}(X \le x) = \int_{-\infty}^x f_X(y)\,\mathrm{d}y$ if X is continuous.
$\mathsf{P}(a < X \le b) = F_X(b) - F_X(a)$.

pmf of random variable $p_X(x) = \mathsf{P}(X = x)$ if X is discrete.

pdf of random variable $f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_X(x)$ if X is continuous.

quantile function $Q_X(p)$ such that $F(Q_X(p)) = p$ if X is continuous. Otherwise, $Q_X(p)$ satisfies $F_X(x) \le p \le F(Q_X(p))$ for $x < Q_X(p)$.

distribution of a function of X $F_Y(y) = \mathsf{P}(h(X) \le y)$ for $Y = h(X)$.

If X is a discrete rv or $h(x)$ takes only countably many values then Y has pmf $p_Y(y) = \mathsf{P}(h(X) = y)$.

If X is a continuous rv and $h(x)$ is a continuous function then Y has pdf $f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} \mathsf{P}(h(X) \le y)$.

binomial theorem $\sum_{k=0}^n \binom{n}{k} a^k b^{n-k} = (a+b)^n$.

geometric sum $\sum_{k=n}^\infty a^k = \frac{a^n}{1-a}$ if $-1 < a < 1$.

exponential expansion $\sum_{k=0}^\infty \frac{a^k}{k!} = \mathrm{e}^a$.

integral of a power function $\int_u^v x^a\,\mathrm{d}x = \frac{v^{a+1} - u^{a+1}}{a+1}$ if $a \ne -1$, and $\int_u^v x^{-1}\,\mathrm{d}x = \log_e(v/u)$.

integral of an exponential function $\int_u^v e^{ax}\,\mathrm{d}x = \frac{1}{a}(\mathrm{e}^{av} - \mathrm{e}^{au})$.

gamma integral $\int_0^\infty x^{a-1}\mathrm{e}^{-x}\,\mathrm{d}x = \Gamma(a) = (a-1)!$ for $a > 0$.

integral of exponential of a quadratic $\int_{-\infty}^\infty \mathrm{e}^{a+bx-cx^2}\,\mathrm{d}x = \sqrt{\frac{\pi}{c}}\,\mathrm{e}^{b^2/(4c)+a}$ for $c > 0$.

discrete uniform(N) pmf $p(x) = \frac{1}{N}$ for $x = 1, 2, \dots, N$.

hypergeometric(N, M, n) pmf $p(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$ for $x = 0, 1, \dots, n$, $M \le N$.

binomial(n, θ) pmf $p(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$ for $x = 0, 1, \dots, n$, $0 < \theta < 1$.

geometric(θ) pmf $p(x) = \theta(1-\theta)^x$ for $x = 0, 1, 2, \dots$, $0 < \theta < 1$.

negative binomial(r, θ) pmf $p(x) = \binom{r-1+x}{r-1}\theta^r(1-\theta)^x$ for $x = 0, 1, 2, \ldots$, $0 < \theta < 1$.

Poisson(λ) pmf $p(x) = \frac{\lambda^x}{x!}\,\mathrm{e}^{-\lambda}$ for $x = 0, 1, 2, \ldots$, $\lambda > 0$.

uniform(a, b) pdf $f(x) = \frac{1}{b-a}$ for $a < x < b$.

exponential(λ) pdf $f(x) = \lambda e^{-\lambda x}$ for $x > 0$, $\lambda > 0$.

gamma(α, λ) pdf $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\,x^{\alpha-1}\mathrm{e}^{-\lambda x}$ for $x > 0$, $\lambda > 0$, $\alpha > 0$.

normal(μ, σ^2) pdf $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}$ for $-\infty < x < \infty$, $\sigma^2 > 0$.

Weibull(α, β) pdf $f(x) = \frac{\alpha}{\beta}\,(x/\beta)^{\alpha-1}\mathrm{e}^{-(x/\beta)^\alpha} I_{(0,\infty)}(x)$ for $\alpha > 0$, $\beta > 0$. $\mathsf{E}(X^k) = \beta^k\,\Gamma(1 + \frac{k}{\alpha})$.

beta(a, b) pdf $f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\,x^{a-1}(1-x)^{b-1}$ for $0 < x < 1$, $a > 0$, $b > 0$.

joint cdf $F_{X,Y}(x, y) = \mathsf{P}(\{X \le x\} \cap \{Y \le y\})$.

joint pmf $p_{X,Y}(x, y) = \mathsf{P}(\{X = x\} \cap \{Y = y\})$, $F_{X,Y}(x, y) = \sum_{u \le x}\sum_{v \le y} p_{X,Y}(u, v)$.

joint pdf $f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$, $F_{X,Y}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u, v)\,\mathrm{d}v\mathrm{d}u$.

marginal pmf/pdf $p_X(x) = \sum_y p_{X,Y}(x, y)$, $p_Y(y) = \sum_x p_{X,Y}(x, y)$;
$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,\mathrm{d}y$, $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,\mathrm{d}x$.

conditional pmf/pdf $p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}$; $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$.

independent random variables $p(x, y) = p_X(x)p_Y(y)$ if (X, Y) is discrete;
$f(x, y) = f_X(x)f_Y(y)$ if (X, Y) is continuous.

discrete convolution $p_{X+Y}(z) = \sum_x p_X(x)p_Y(z - x)$ for independent X, Y.

continuous convolution $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z - x)\,\mathrm{d}x$ for independent X, Y.

cdf of minimum $F_{\min(X_1,\ldots,X_n)}(u) = 1 - (1 - F_{X_1}(u)) \times \cdots \times (1 - F_{X_n}(u))$ for independent $X_1, \ldots, X_n$.

cdf of maximum $F_{\max(X_1,\ldots,X_n)}(u) = F_{X_1}(u) \times \cdots \times F_{X_n}(u)$ for independent $X_1, \ldots, X_n$.

expectation for a discrete rv $\mathsf{E}(h(X)) = \sum_x h(x)p_X(x)$.

expectation for a continuous rv $\mathsf{E}(h(X)) = \int_{-\infty}^{\infty} h(x)f_X(x)\,\mathrm{d}x$.

mean and variance $\mu_X = \mathsf{E}(X)$; $\sigma_X^2 = \mathsf{Var}(X) = \mathsf{E}((X - \mu_X)^2) = \mathsf{E}(X^2) - \mu_X^2$.

standard deviation $\sigma_X = \sqrt{\mathsf{Var}(X)}$.

covariance and correlation $\mathsf{Cov}(X, Y) = \mathsf{E}((X - \mu_X)(Y - \mu_Y)) = \mathsf{E}(XY) - \mu_X\mu_Y$; $\mathsf{Corr}(X, Y) = \frac{\mathsf{Cov}(X,Y)}{\sigma_X\sigma_Y}$.
For independent X and Y, $\mathsf{Cov}(X, Y) = \mathsf{Corr}(X, Y) = 0$.

expectation of a sum $\mathsf{E}(a_1X_1 + \cdots + a_nX_n) = a_1\mathsf{E}(X_1) + \cdots + a_n\mathsf{E}(X_n)$.

expectation of a product If $X_1, \ldots, X_n$ are independent, $\mathsf{E}\left(\prod_{i=1}^n h_i(X_i)\right) = \prod_{i=1}^n \mathsf{E}(h_i(X_i))$.

variance of a sum $\mathsf{Var}(aX + bY) = a^2\,\mathsf{Var}(X) + 2ab\,\mathsf{Cov}(X, Y) + b^2\,\mathsf{Var}(Y)$.

variance of a sum of independent rvs $\mathsf{Var}(a_1X_1 + \cdots + a_nX_n) = a_1^2\,\mathsf{Var}(X_1) + \cdots + a_n^2\,\mathsf{Var}(X_n)$.

moments k-th moment is $\mathsf{E}(X^k)$, $k = 1, 2, \ldots$