**Statistics 630 – Exam I**
**Wednesday, 6 October 2021**

## INSTRUCTIONS:

1. You have 50 minutes to complete the exam.
2. There are 5 pages including this cover sheet and the formula sheets.
3. Questions 1–4 are multiple choice and worth 5 points each. Questions 5–8 require solutions to be worked out and are 10 points each. (60 points total.)
4. Please write out your answers *in order* on blank sheets of paper, explaining your steps. You may refer to theorems by name/description rather than by its number in the book.
5. 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.
6. You may use but mostly do not need 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.
7. Do not discuss or provide any information to anyone concerning any of the questions on this exam until your papers are returned or I post my solutions.

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

1. (5 points) If $x_p$ is the $p$-quantile for $X$ then the $p$-quantile for $Y = aX + b$, $a > 0$, is

   (a) $x_{ap+b}$.
   (b) $ax_p + b$.
   (c) $x_{a^2p}$.
   (d) $a^2x_p$.
   (e) not known without the distribution for $Y$.

2. (5 points) $(W, Z)$ has a bivariate normal pdf which can be expressed in the following ways:

$$f(w,z) = \frac{e^{-(w^2-2wz+2z^2)/2}}{2\pi} = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \times \frac{e^{-(w-z)^2/2}}{\sqrt{2\pi}}.$$

   Based on this we recognize that

   (a) $Z \sim$ normal$(0,1)$ and $W \sim$ normal$(z,1)$, independent.
   (b) $Z \sim$ Weibull$(2,\sqrt{2})$ and $W \sim$ normal$(z,1)$, independent.
   (c) $Z \sim$ normal$(0,1)$ and $W$, given $Z = z$, has conditional normal$(z,1)$ distribution.
   (d) $Z \sim$ Weibull$(2,\sqrt{2})$ and $W$, given $Z = z$, has conditional normal$(z,1)$ distribution.
   (e) $W \sim$ normal$(0,1)$ and $Z$, given $W = w$, has conditional normal$(w,1)$ distribution.

3. (5 points) An experiment will be conducted which consists of going to each of three schools and randomly sampling 10 different students. The schools have 160, 140 and 100 students, respectively. Investigators will observe how many students in the samples have COVID-19 antibodies. The size of the complete sample space is

   (a) $10^3$.
   (b) $\binom{400}{30}$.
   (c) $\binom{160}{10} + \binom{140}{10} + \binom{100}{10}$.
   (d) $\binom{160}{10} \times \binom{140}{10} \times \binom{100}{10}$.
   (e) $160^{10} \times 140^{10} \times 100^{10}$.

4. (5 points) $S$ and $T$ have the following joint density function: $f(s,t) = 120st(1-s-t)$ for $s \geq 0$, $t \geq 0$, $s + t \leq 1$. The marginal density function for $S$ is (by inspection: little or no computation needed)

   (a) $20s(1-s)^3$ for $0 \leq s \leq 1$.
   (b) $12s(1-t-s)^2$ for $0 \leq s \leq 1-t$.
   (c) $20s - 60s^2$ for $0 \leq s \leq 1$.
   (d) $20s - 60s^2$ for $0 \leq s \leq 1-t$.
   (e) $60s(1-s-t)$ for $0 \leq s \leq 1$.

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

5. (10 points) Randy has 2 black mice and 7 white mice in his hat. He selects one at random. If the mouse is black, he rolls two fair dice and gets the total. If the mouse is white, he rolls three fair dice and gets the total.

   (a) What is the probability that the total is 4?

   (b) Randy tells Sally that the total is 4 but not which color of mouse he selected. Help Sally determine the probability that the mouse was black given the information she has.

6. (10 points) $X \sim \text{exponential}(1)$ with pdf $f_X(t) = e^{-t}$, $t \geq 0$, and $Y \sim \text{gamma}(2, 1)$ with pdf $f_Y(t) = te^{-t}$, $t \geq 0$. Show that their cumulative distribution functions satisfy $F_X(t) \geq F_Y(t)$ for all real $t$. [This makes $Y$ *stochastically larger* than $X$.]

7. The number of fire ant hills in my yard at a given time is a Poisson(7) random variable. The number in my neighbor's yard is independent of mine and has Poisson(3) distribution. What is the probability that I have at least 4 fire ant hills and my neighbor does not?

8. (10 points) Suppose positive random variable $R$ has cdf $F_R(r) = \mathsf{P}(R \leq r) = 1 - e^{-r^2/4}$, $r > 0$. Let $X = R^{2/3}$ and determine the cdf and pdf for $X$.

## Formulas for Exam I

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

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

**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 $B_1, \ldots, B_n$ are disjoint, $\bigcup_{k=1}^{n} B_k = \Omega$.

**Bayes' rule** $\mathsf{P}(B_j \mid A) = \dfrac{\mathsf{P}(A|B_j)\mathsf{P}(B_j)}{\sum_{k=1}^{n} \mathsf{P}(A|B_k)P(B_k)}$ if $B_1, \ldots, B_n$ 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** $\int_u^v e^{ax}\,\mathrm{d}x = \frac{1}{a}\left(\mathrm{e}^{av} - \mathrm{e}^{au}\right)$.

**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 $x = 1, 2, \ldots, N$.

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

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

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

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

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

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

**normal**$(\mu,\sigma^2)$ **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**$(\alpha,\beta)$ **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 $X_1, \ldots, X_n$.