TEST #1
STA 5326
September 30, 2020

Please read the following directions. (They are included in the directions I posted to the Canvas Announcements, so you can skip the following if you have already read those directions.)

- The exam is closed book and closed notes. No books, notes, or internet resources are allowed.
- A copy of the Table of Common Distributions from the back of our textbook is attached to the end of your exam.
- During the exam, you need only a supply of blank paper and writing implements (pens, pencils, erasers, etc). If you wish, you may use an ordinary scientific calculator (TI-86 or below is fine).
- Partial credit is available. (If you know part of a solution, write it down. If you know an approach to a problem, but cannot carry it out – write down this approach. If you know a useful result, write it down.)
- You must show and explain your work (including your calculations). **No credit is given without work or explanation!** But don't get carried away! Give enough explanation and work so that what you have done is clearly understandable.
- Make sure that the grader can easily see how you get from one step to the next. If you needed scratch paper to work something out, this work should be copied to your written solutions.
- **If your answer is valid only for a certain range of values, this should be stated as part of your answer. For example, if a density is zero outside of some interval, this interval should be stated explicitly.**
- You should give only one answer to each problem. **Circle your answer** if there is any chance for confusion.
- Simplify your answers when it is easy to do so. But more difficult arithmetic does **not** have to be done completely. Answers can be left as fractions or products. You do not have to evaluate large binomial coefficients, factorials or powers. Answers can be left as summations (unless there is a simple closed form such as when summing a geometric or exponential series).
- All algebra and calculus must be done completely. (Only arithmetic can be left incomplete.)
- Do **not** quote homework results. If you wish to use a result from homework in a solution, you must prove this result.
- All the work on the exam should be your own. No "cooperation" is allowed.
- The exam has **6** problems and there are a total of **100** points.

Problem 1. A sample of four balls is chosen at random (without replacement) from an urn containing 9 red balls and 6 green balls. Let X be the number of red balls in the sample. Answer the following.

(a) (8%) Find $P(X = 2)$.

(b) (8%) Define $Y = 1/(X+1)$ and let $F_Y(y)$ denote the cdf (cumulative distribution function) of Y. Find $F_Y(0.444)$.

Problem 2. (10%) Let X be a continuous random variable with pdf $f(x)$ and cdf $F(x)$. Assume the pdf f satisfies $f(x) > 0$ for $-\infty < x < \infty$. For fixed numbers a and b with $a < b$, define the function

$$g(x) = \begin{cases} \dfrac{f(x)}{F(b) - F(a)} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

Prove that $g(x)$ is a pdf. (Give a detailed argument verifying all the requirements for g to be a pdf.)

Problem 3. A newly married couple agree that they would like to have both a boy and a girl child but cannot afford more than three children. So they plan to have two children and stop at that point if they have both a boy and a girl, but will go on and have a third child if they do not. Suppose the couple follow this plan. (Assume that the sex of different children is independent and that boys and girls have equal probability 1/2. Also assume twins have probability zero.)

(a) (8%) Conditional on the couple having at least one boy, what is the probability all their children are boys?

(b) (8%) The couple later wins the lottery and now can afford to have as many children as they want. So they change their plan and decide to keep having children until they have both a boy and a girl (and then they will stop). What is the probability they will have **more than** 5 children?

Problem 4. Let X have pdf

$$f_X(x) = (x+4)/18 \quad \text{for } -4 < x < 2$$

and define $Y = g(X)$ where the function $g(x)$ is given by

$$g(x) = \begin{cases} 2x^2 & \text{for } x \geq 0 \\ x^2 & \text{for } x < 0\,. \end{cases}$$

(a) (6%) Find $F_X(x)$, the cdf of X.

(b) (6%) Let A be the interval $A = (2, 8) = \{y : 2 < y < 8\}$. Find $g^{-1}(A)$.

(c) (6%) Use your answers to the previous parts to find $P_Y(A)$, where P_Y is the induced probability function of Y.

(d) (8%) Find $f_Y(y)$, the pdf of Y.

Problem 5. Let

$$Y = 2\sum_{i=1}^{\infty} \frac{Z_i}{3^i}$$

where $Z_1, Z_2, Z_3, \ldots$ are independent with $P(Z_i = 0) = P(Z_i = 1) = 1/2$.

(a) (8%) Find EY.

(b) (8%) Find $\text{Var}(Y)$.

Problem 6. (16%) A fair dodecahedron (a 12-sided solid) has its sides labeled 1, 2, ..., 12. The dodecahedron is rolled five times. Let X_1, X_2, X_3, X_4, X_5 denote the observed sequence of values. Find the value of the conditional probability

$$P(X_1 < X_2 < X_3 < X_4 < X_5 \,|\, X_1 \le X_2 \le X_3 \le X_4 \le X_5)$$

Table of Common Distributions

Do Not Mark !!

Discrete Distributions

Bernoulli(p)

pmf	$P(X = x\|p) = p^x(1-p)^{1-x}; \quad x = 0, 1; \quad 0 \le p \le 1$
mean and variance	$\mathrm{E}X = p, \quad \mathrm{Var}\, X = p(1-p)$
mgf	$M_X(t) = (1-p) + pe^t$

Binomial(n, p)

pmf	$P(X = x\|n, p) = \binom{n}{x} p^x(1-p)^{n-x}; \quad x = 0, 1, 2, \ldots, n; \quad 0 \le p \le 1$
mean and variance	$\mathrm{E}X = np, \quad \mathrm{Var}\, X = np(1-p)$
mgf	$M_X(t) = [pe^t + (1-p)]^n$

notes Related to Binomial Theorem (Theorem 3.1.1). The *multinomial* distribution (Definition 4.6.1) is a multivariate version of the binomial distribution.

Discrete Uniform

pmf	$P(X = x\|N) = \frac{1}{N}; \quad x = 1, 2, \ldots, N; \quad N = 1, 2, \ldots$
mean and variance	$\mathrm{E}X = \frac{N+1}{2}, \quad \mathrm{Var}\, X = \frac{(N+1)(N-1)}{12}$
mgf	$M_X(t) = \frac{1}{N} \sum_{i=1}^{N} e^{it}$

Geometric(p)

pmf $P(X=x|p)=p(1-p)^{x-1}; \quad x=1,2,\ldots; \quad 0\le p\le 1$

mean and variance $\mathrm{E}X=\frac{1}{p}, \quad \mathrm{Var}\,X=\frac{1-p}{p^2}$

mgf $M_X(t)=\frac{pe^t}{1-(1-p)e^t}, \quad t<-\log(1-p)$

notes $Y=X-1$ is negative binomial$(1,p)$. The distribution is *memoryless*: $P(X>s|X>t)=P(X>s-t)$.

Hypergeometric

pmf $P(X=x|N,M,K)=\frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}; \quad x=0,1,2,\ldots,K;$
$M-(N-K)\le x\le M; \quad N,M,K\ge 0$

mean and variance $\mathrm{E}X=\frac{KM}{N}, \quad \mathrm{Var}\,X=\frac{KM}{N}\frac{(N-M)(N-K)}{N(N-1)}$

notes If $K<<M$ and N, the range $x=0,1,2,\ldots,K$ will be appropriate.

Negative binomial(r,p)

pmf $P(X=x|r,p)=\binom{r+x-1}{x}p^r(1-p)^x; \quad x=0,1,\cdots; \quad 0\le p\le 1$

mean and variance $\mathrm{E}X=\frac{r(1-p)}{p}, \quad \mathrm{Var}\,X=\frac{r(1-p)}{p^2}$

mgf $M_X(t)=\left(\frac{p}{1-(1-p)e^t}\right)^r, \quad t<-\log(1-p)$

notes An alternate form of the pmf is given by $P(Y=y|r,p)=\binom{y-1}{r-1}p^r(1-p)^{y-r}$, $y=r,r+1,\ldots$. The random variable $Y=X+r$. The negative binomial can be derived as a gamma mixture of Poissons. (See Exercise 4.34.)

Poisson(λ)

pmf $P(X=x|\lambda)=\frac{e^{-\lambda}\lambda^x}{x!}; \quad x=0,1,\ldots; \quad 0\le\lambda<\infty$

mean and variance $\mathrm{E}X=\lambda, \quad \mathrm{Var}\,X=\lambda$

mgf $M_X(t)=e^{\lambda(e^t-1)}$

Continuous Distributions

Beta(α, β)

pdf $f(x|\alpha, \beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 \le x \le 1, \quad \alpha > 0, \quad \beta > 0$

mean and variance $\mathrm{E}X = \frac{\alpha}{\alpha+\beta}, \quad \mathrm{Var}\, X = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

mgf $M_X(t) = 1 + \sum_{k=1}^{\infty} \left(\prod_{r=0}^{k-1} \frac{\alpha+r}{\alpha+\beta+r} \right) \frac{t^k}{k!}$

notes The constant in the beta pdf can be defined in terms of gamma functions, $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. Equation (3.2.18) gives a general expression for the moments.

Cauchy(θ, σ)

pdf $f(x|\theta, \sigma) = \frac{1}{\pi\sigma} \frac{1}{1+\left(\frac{x-\theta}{\sigma}\right)^2}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty, \quad \sigma > 0$

mean and variance do not exist

mgf does not exist

notes Special case of Student's t, when degrees of freedom $= 1$. Also, if X and Y are independent $\mathrm{n}(0, 1)$, X/Y is Cauchy.

Chi squared

pdf $f(x|p) = \frac{1}{\Gamma(p/2)2^{p/2}} x^{(p/2)-1} e^{-x/2}; \quad 0 \le x < \infty; \quad p = 1, 2, \ldots$

mean and variance $\mathrm{E}X = p, \quad \mathrm{Var}\, X = 2p$

mgf $M_X(t) = \left(\frac{1}{1-2t} \right)^{p/2}, \quad t < \frac{1}{2}$

notes Special case of the gamma distribution.

Double exponential(μ, σ)

pdf $f(x|\mu, \sigma) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$

mean and variance $\mathrm{E}X = \mu, \quad \mathrm{Var}\, X = 2\sigma^2$

mgf $M_X(t) = \frac{e^{\mu t}}{1-(\sigma t)^2}, \quad |t| < \frac{1}{\sigma}$

notes Also known as the *Laplace* distribution.

Exponential(β)

pdf $f(x|\beta) = \frac{1}{\beta}e^{-x/\beta}, \quad 0 \le x < \infty, \quad \beta > 0$

mean and variance $\mathrm{E}X = \beta, \quad \mathrm{Var}\, X = \beta^2$

mgf $M_X(t) = \frac{1}{1-\beta t}, \quad t < \frac{1}{\beta}$

notes Special case of the gamma distribution. Has the *memoryless* property. Has many special cases: $Y = X^{1/\gamma}$ is *Weibull*, $Y = \sqrt{2X/\beta}$ is *Rayleigh*, $Y = \alpha - \gamma \log(X/\beta)$ is *Gumbel*.

F

pdf $$f(x|\nu_1,\nu_2) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2}\frac{x^{(\nu_1-2)/2}}{\left(1+\left(\frac{\nu_1}{\nu_2}\right)x\right)^{(\nu_1+\nu_2)/2}}; \quad 0 \le x < \infty;$$

$\nu_1, \nu_2 = 1, \ldots$

mean and variance $\mathrm{E}X = \frac{\nu_2}{\nu_2-2}, \quad \nu_2 > 2,$

$$\mathrm{Var}\, X = 2\left(\frac{\nu_2}{\nu_2-2}\right)^2 \frac{(\nu_1+\nu_2-2)}{\nu_1(\nu_2-4)}, \quad \nu_2 > 4$$

moments (mgf does not exist) $$\mathrm{E}X^n = \frac{\Gamma\left(\frac{\nu_1+2n}{2}\right)\Gamma\left(\frac{\nu_2-2n}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_2}{\nu_1}\right)^n, \quad n < \frac{\nu_2}{2}$$

notes Related to chi squared ($F_{\nu_1,\nu_2} = \left(\frac{\chi^2_{\nu_1}}{\nu_1}\right) / \left(\frac{\chi^2_{\nu_2}}{\nu_2}\right)$, where the χ^2s are independent) and t ($F_{1,\nu} = t^2_\nu$).

Gamma(α, β)

pdf $f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}, \quad 0 \le x < \infty, \quad \alpha,\ \beta > 0$

mean and variance $\mathrm{E}X = \alpha\beta, \quad \mathrm{Var}\, X = \alpha\beta^2$

mgf $M_X(t) = \left(\frac{1}{1-\beta t}\right)^\alpha, \quad t < \frac{1}{\beta}$

notes Some special cases are exponential ($\alpha = 1$) and chi squared ($\alpha = p/2, \beta = 2$). If $\alpha = \frac{3}{2}$, $Y = \sqrt{X/\beta}$ is *Maxwell*. $Y = 1/X$ has the *inverted gamma distribution*. Can also be related to the Poisson (Example 3.2.1).

Logistic(μ, β)

pdf $f(x|\mu,\beta) = \frac{1}{\beta}\frac{e^{-(x-\mu)/\beta}}{[1+e^{-(x-\mu)/\beta}]^2}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty,$
$\beta > 0$

mean and variance $EX = \mu, \quad \text{Var}\, X = \frac{\pi^2\beta^2}{3}$

mgf $M_X(t) = e^{\mu t}\Gamma(1-\beta t)\Gamma(1+\beta t), \quad |t| < \frac{1}{\beta}$

notes The cdf is given by $F(x|\mu,\beta) = \frac{1}{1+e^{-(x-\mu)/\beta}}$.

Lognormal(μ, σ^2)

pdf $f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\frac{e^{-(\log x-\mu)^2/(2\sigma^2)}}{x}, \quad 0 \le x < \infty, \quad -\infty < \mu < \infty,$
$\sigma > 0$

mean and variance $EX = e^{\mu+(\sigma^2/2)}, \quad \text{Var}\, X = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2}$

moments (mgf does not exist) $EX^n = e^{n\mu+n^2\sigma^2/2}$

notes Example 2.3.5 gives another distribution with the same moments.

Normal(μ, σ^2)

pdf $f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty,$
$-\infty < \mu < \infty, \quad \sigma > 0$

mean and variance $EX = \mu, \text{Var}\, X = \sigma^2$

mgf $M_X(t) = e^{\mu t+\sigma^2 t^2/2}$

notes Sometimes called the *Gaussian* distribution.

Pareto(α, β)

pdf $f(x|\alpha,\beta) = \frac{\beta\alpha^\beta}{x^{\beta+1}}, \quad \alpha < x < \infty, \quad \alpha > 0, \quad \beta > 0$

mean and variance $EX = \frac{\beta\alpha}{\beta-1}, \quad \beta > 1,$
$\text{Var}\, X = \frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)}, \quad \beta > 2$

mgf does not exist

t

pdf $f(x|\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \frac{1}{\left(1+\left(\frac{x^2}{\nu}\right)\right)^{(\nu+1)/2}}, \quad -\infty < x < \infty, \quad \nu = 1, \ldots$

mean and variance $\mathrm{E}X = 0, \quad \nu > 1,$

$\mathrm{Var}\, X = \frac{\nu}{\nu-2}, \quad \nu > 2$

moments (mgf does not exist) $\mathrm{E}X^n = \frac{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{\nu-n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)} \nu^{n/2}$ if $n < \nu$ and even,

$\mathrm{E}X^n = 0$ if $n < \nu$ and odd.

notes Related to F ($F_{1,\nu} = t_\nu^2$).

Uniform(a, b)

pdf $f(x|a, b) = \frac{1}{b-a}, \quad a \le x \le b$

mean and variance $\mathrm{E}X = \frac{b+a}{2}, \quad \mathrm{Var}\, X = \frac{(b-a)^2}{12}$

mgf $M_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}$

notes If $a = 0$ and $b = 1$, this is a special case of the beta ($\alpha = \beta = 1$).

Weibull(γ, β)

pdf $f(x|\gamma, \beta) = \frac{\gamma}{\beta} x^{\gamma-1} e^{-x^\gamma/\beta}, \quad 0 \le x < \infty, \quad \gamma > 0, \quad \beta > 0$

mean and variance $\mathrm{E}X = \beta^{1/\gamma}\Gamma\left(1 + \frac{1}{\gamma}\right), \quad \mathrm{Var}\, X = \beta^{2/\gamma}\left[\Gamma\left(1 + \frac{2}{\gamma}\right) - \Gamma^2\left(1 + \frac{1}{\gamma}\right)\right]$

moments $\mathrm{E}X^n = \beta^{n/\gamma}(1 + \frac{n}{\gamma})$

notes The mgf exists only for $\gamma \ge 1$. Its form is not very useful. A special case is exponential ($\gamma = 1$).