Please do not write in the boxes immediately below.

problem	1	2	3	4	5	6	EC	total
points								

STAT 232 Categorical Data Analysis Spring 2025 Exam 1

February 26, 2025

Your name____

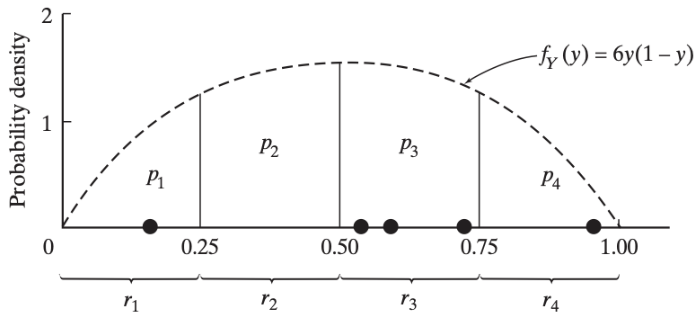
The exam has 7 different printed sides of exam problems and 1 side workspace.

Duration of the exam is 90 minutes. There are 6 problems, worth 20 points each. From Problems 1 – 6, only 5 problems will be graded. If you solve all Problems 1 – 6, you must cross out the problem in the box above that must not be graded. If you solve all Problems 1 – 6 and do not cross out a problem, only the first five problems will be graded. Show all your work for full credit. Books, notes etc. are prohibited. Calculators, cellphones, earphones, AirPods and cheat sheets are NOT permitted.

1. (a) Five observations are drawn at random from the pdf

$$f_Y(y) = 6y(1-y), \quad 0 \le y \le 1.$$

What is the probability that one of the observations lies in the interval $[0, 0.25)$, none in the interval $[0.25, 0.50)$, three in the interval $[0.50, 0.75)$, and one in the interval $[0.75, 1.00]$?



(b) In the daily production of a certain kind of rope, the number of defects per foot Y is assumed to have a Poisson distribution with mean $\lambda = 2$. The profit per foot when the rope is sold is given by X, where $X = 50 - 2Y - Y^2$. Find the expected profit per foot.

2. (a) Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from an exponential distribution with unknown parameter θ whose pdf is given by

$$f_Y(y;\theta) = \frac{1}{\theta}\, e^{-y/\theta}, \quad 0 \leq y < \infty;\, 0 < \theta < \infty.$$

Find the Maximum Likelihood Estimator for θ.

(b) Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from the exponential pdf given above. Find the form of the likelihood ratio test statistic for $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

3. (a) Let $X_1, X_2, \ldots, X_n$ be a random sample from a discrete pdf $p_X(k;\theta)$, where $\theta = E(X)$ is an unknown parameter. Consider the estimator

$$\hat{\theta} = \sum_{i=1}^{n} a_i X_i,$$

Where the a_i's are constants. For what values of $a_1, a_2, \ldots, a_n$ will $\hat{\theta}$ be unbiased?

(b) A random sample of size 2, Y_1 and Y_2, is drawn from the pdf

$$f_Y(y;\theta) = 3y\theta^2, \;\; 0 < y < \frac{1}{\theta}.$$

What must c equal if the statistic $c(Y_1 + 3Y_2)$ is to be an unbiased estimator for $\frac{1}{\theta}$?

4. (a) A random sample of 500 measurements on the length of stay in hospitals had sample mean 5.4 days and sample standard deviation 3.1 days. A federal regulatory agency hypothesizes that the average length of stay is in excess of 5 days. Do the data support this hypothesis? Use $\alpha = .01$. **Hint:** $z_{0.01} = 2.33$, $z_{0.005} = 2.575$

(b) **Windows or Mac?** A recent survey of 85 Holy Cross students indicated that 68 have an Apple laptop / iPad. We are looking to test the claim that 2/3 of students use an Apple product.

i. State an appropriate null and alternative hypothesis.

ii. Conduct a Wald test using the χ^2 distribution to compute a p-value. At the 2% level of significance, is there sufficient evidence to conclude that the proportion of students with an Apple product is not 2/3? **Hint:** $P(\chi^2 > 9.440) = 0.00212$.

iii. Perform the same analysis as in part (b), but this time using a score test. **Hint:** $P(\chi^2 > 6.797) = 0.00913$

5. Two detergents were tested for their ability to remove stains of a certain type. An inspector judged the first one to be successful on 63 out of 91 independent trials and the second one to be successful on 42 out of 79 independent trials. **Hint:** $z_{0.1} = 1.2816$, $z_{0.05} = 1.6449$.

   (a) Do the two detergents appear to differ with respect to their ability to remove stains, at the 10% significance level?

   (b) Construct a 90% confidence interval for the difference in proportions $p_1 - p_2$.

   (c) Is the value $p_1 - p_2$ inside or outside this interval?

   (d) Based on the interval, should the null hypothesis be rejected? Why?

   (e) How does the conclusion that you reached compare with your conclusion in part (a)?

6. Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a Poisson distribution with density $p_X(k;\lambda) = \dfrac{e^{-\lambda}\,\lambda^k}{k!}$. Use the conjugate gamma (α, β) prior for λ to find the posterior density for λ. Also, find the Bayes estimator for λ.

WORKSHEET