MATH 401 Fall 2014

Homework Assignments:

Homework 1, Due September 10

1) If A and B are two sets, draw Venn diagrams to verify the following:

   a) $A = (A \cap B) \cup (A \cap \overline{B})$.

   b) If $B \subset A$ then $A = B \cup (A \cap \overline{B})$.

2) A vehicle arriving at an intersection can turn right, turn left, or continue straight ahead. The experiment consists of observing the movement of a single vehicle through the intersection.

   a) List the sample space for this experiment.

   b) Assuming that all sample points are equally likely, find the probability that the vehicle turns.

3) Suppose two balanced coins are tossed and the upper faces are observed.

   a) List the sample points for this experiment.

   b) Assign a reasonable probability to each sample point. (Are the sample points equally likely?)

   c) Let A denote the event that *exactly* one head is observed and B the event that *at least* one head is observed. List the sample points in both A and B.

   d) From your answer to part (c), find $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$, and $P(\overline{A} \cup B)$.

4) If A and B are events and $B \subset A$, use the result derived in (1b) above and the probability axioms to prove that $P(A) = P(B) + P(A \cap \overline{B})$.

5) According to *Webster's New Collegiate Dictionary*, a divining rod is "a forked rod believed to indicate [divine] the presence of water or minerals by dipping downward when held over a vein." To test the claims of a divining rod expert, skeptics bury four cans in the ground, two empty and two filled with water. The expert is led to the four cans and told that two contain water. He uses the divining rod to test each of the four cans and decide which two contain water.

   a) List the sample space for this experiment.

   b) If the divining rod is completely useless for locating water, what is the probability that the expert will correctly identify (by guessing) both of the cans containing water?

Homework 2, Due September 17

1) A group of three undergraduates and five graduate students are available to fill certain student government posts. If four students are to be randomly selected from this group, find the probability that exactly two undergraduates will be among the four chosen.

2) A survey of consumers in a particular community showed that 10% were dissatisfied with plumbing jobs done in their homes. Half the complaints dealt with plumber A, who does 40% of the plumbing jobs in the town. Find the probability that a consumer will obtain:

   a) an unsatisfactory plumbing job, given that the plumber was A.

   b) a satisfactory plumbing job, given that the plumber was A.

3) Suppose that A and B are two events such that $P(A) = .6$ and $P(B) = .3$.

   a) Is it possible that $P(A \cap B) = .1$? Why or why not?

   b) What is the smallest possible value for $P(A \cap B)$?

   c) Is it possible that $P(A \cap B) = .7$? Why or why not?

   d) What is the largest possible value for $P(A \cap B)$?

4) If A and B are two events, prove that $P(A \cap B) \geq 1 - P(\overline{A}) - P(\overline{B})$. [Note: This is a simplified version of the *Bonferroni inequality*.]

5) Persons entering a blood bank are such that 1 in 3 have type O^+ blood and 1 in 15 have type O^- blood. Consider three randomly selected donors for the blood bank. Let X denote the number of donors with type O^+ blood and Y denote the number with type O^- blood. Find the probability distributions for X and Y. Also find the probability distribution for $X + Y$, the number of donors who have type O blood.

Homework 3, Due September 24

1) Let Y be a random variable with $p(y)$ given in the accompanying table. Find $E(Y), E\left(\frac{1}{Y}\right), E(Y^2 - 1)$, and $V(Y)$.

y	1	2	3	4
$p(y)$	.4	.3	.2	.1

2) The manufacturer of a low-calorie dairy drink wishes to compare the taste appeal of a new formula (formula B with that of the standard formula (formula A). Each of four judges is given three glasses in random order, two containing formula A and the other containing formula B. Each judge is asked to state which glass he or she most enjoyed. Suppose that the two formulas are equally attractive. Let Y be the number of judges stating a preference for the new formula.

a) Find the probability function for Y.

b) What is the probability that at least three of the four judges state a preference for the new formula?

c) Find the expected value and variance of Y.

3) If Y has a geometric distribution with success probability .3, what is the largest value, y_0, such that $P(Y > y_0) \geq .1$?

4) The employees of a firm that manufactures insulation are being tested for indications of asbestos in their lungs. The firm is requested to send three employees who have positive indications of asbestos on to a medical center for further testing. If 40% of the employees have positive indications of asbestos in their lungs, find the probability that ten employees must be tested in order to find three positives.

5) A shipment of 20 cameras includes 3 that are defective. What is the minimum number of cameras that must be selected if we require that $P(\text{at least 1 defective}) \geq .8$?

Homework 4, Due October 1

1) By finding derivatives at $t = 0$, give the mean and variance of the random variables associated with the following moment generating functions.

a) $m(t) = \left(\frac{1}{3}e^t + \frac{2}{3}\right)^5$.

b) $m(t) = \frac{e^t}{2-e^t}$.

c) $m(t) = e^{2(e^t-1)}$.

2) If Y is a random variable with moment-generating function $m(t)$ and if W is given by $W = aY + b$, show that the moment-generating function of W is $m_W(t) = e^{tb}m(at)$.

3) Let Y have the density function given by $f(y) = \begin{cases} .2 & -1 < y \leq 0 \\ .2 + cy & 0 < y \leq 1 \\ 0 & \text{elsewhere} \end{cases}$.

a) Find c.

b) Derive $F(y)$.

c) Graph $f(y)$and $F(y)$.

d) Using $F(y)$ find $F(-1), F(0), F(1), P(0 \leq Y \leq .5)$, and $P(Y > .5|Y > .1)$.

e) Find the mean and variance of Y.

4) The median of the distribution of a continuous random variable Y is the value $\phi_{.5}$ where $F(\phi_{.5}) = .5$. Find the median for the uniform distribution over the interval (θ_1, θ_2).

5) Show that the maximum value of the normal density with parameters μ and σ is $\frac{1}{\sigma\sqrt{2\pi}}$ and occurs when $y = \mu$.

Homework 5, Due October 13

1) Three balanced coins are tossed independently. One of the variables of interest is X, the number of heads. The other variable of interest is Y, the amount of money won on a side bet in the following manner. If the first head occurs on the first toss, you win \$1. If the first head occurs on toss 2 or on toss 3 you win \$2 or \$3, respectively. If no heads appear, you lose \$1 (that is, win $-$\$1).

   a) Find the joint probability function for X and Y.

   b) What is the probability that fewer than 3 heads will occur and you will win \$1 or less? [That is, find $F(2,1)$.]

2) Suppose that the random variables X and Y have joint probability density function $f(x, y)$ given by $f(x,y) = \begin{cases} 6x^2y & 0 < x < y, x + y \leq 2 \\ 0 & \text{elsewhere} \end{cases}$.

   a) Verify that this is a valid joint density function.

   b) What is the probability that $X + Y$ is less than 1?

   c) Show that X and Y are *dependent* random variables.

3) Using the joint probability density function $f(x,y) = \begin{cases} 4xy & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{elsewhere} \end{cases}$, find:

   a) the marginal density functions for X and Y.

   b) $P\left(X \leq \frac{1}{2} \middle| Y \geq \frac{3}{4}\right)$.

   c) the conditional density function of X given $Y = y$.

   d) the conditional density function of Y given $X = x$.

   e) $P\left(X \leq \frac{3}{4} \middle| Y = \frac{1}{2}\right)$.

   f) Expected Value and Variance of X.

   g) $E(X - Y)$.

Homework 6, Due October 20

1) The joint probability distribution of X, the number of married executives, and Y, the number of never-married executives, is given by $p(x,y) = \frac{\binom{4}{x}\binom{3}{y}\binom{2}{3-x-y}}{\binom{9}{3}}$, where x and y are integers between 0 and 3, and their sum is between 1 and 3. Find $\mathrm{Cov}(X,Y)$.

2) Let Z be a standard normal random variable and let $X = Z$ and $Y = Z^2$.

   a) What are $E(X)$ and $E(Y)$?

   b) What is $E(XY)$?

   c) What is $\mathrm{Cov}(X,Y)$?

   d) Notice that $P(Y > 1 \mid X > 1) = 1$. Are X and Y independent?

3) Life lengths for two components I and II are X and Y, respectively. The joint density is $f(x,y) = \begin{cases} \left(\frac{y}{8}\right) e^{\frac{-(x+y)}{2}} & 0 \le x, 0 \le y \\ 0 & \text{elsewhere} \end{cases}$.

   The cost C of replacing the two components depends upon their length of life at failure and is given by $C = 50 + 2Y + 4Y$. Find $E(C)$ and $V(C)$.

4) Let Y be a random variable with a density function given by $f(y) = \begin{cases} \left(\frac{3}{2}\right) y^2 & -1 \le y \le 1 \\ 0 & \text{elsewhere} \end{cases}$.

   a) Find the density function of $U = 3Y$.

   b) Find the density function of $V = 3 - Y$.

   c) Find the density function of $W = Y^2$.

5) The joint distribution of amount of pollutant emitted from a smokestack without a cleaning device (X) and a similar smokestack with a cleaning device (Y) is $f(x,y) = \begin{cases} 1 & 0 \le x \le 2, 0 \le y \le 1, 2y \le x \\ 0 & \text{elsewhere} \end{cases}$.

   The reduction in amount of pollutant due to the cleaning device is given by $U = X - Y$.

   a) Find the probability density function for U.

   b) Find $E(U)$ using the probability density function.

Homework 7, Due October 27

1) The Weibull density function is given by $f(y) = \begin{cases} \frac{1}{\alpha} m y^{m-1} e^{\frac{-y^m}{\alpha}} & y > 0 \\ 0 & \text{elsewhere} \end{cases}$, where α and m are positive constants.

   This density function is often used as a model for the lengths of life of physical systems. Suppose Y has the Weibull density just given.

   a) Find the probability density function of $U = Y^m$.

   b) Find $E(Y^k)$ for any positive integer k.

2) The Rayleigh density function is given by $f(y) = \begin{cases} \frac{2y}{\theta} e^{\frac{-y^2}{\theta}} & y > 0 \\ 0 & \text{elsewhere} \end{cases}$, where θ is a positive constant.

   This density function is sometimes used by engineers as a model for the lengths of life of electronic components. Suppose Y has the Rayleigh density just given.

   a) Find the probability density function of $U = Y^2$.

   b) Find $E(Y)$ and $V(Y)$, using the result in part a).

3) Let X and Y be independent Poisson random variables with means λ_x and λ_y, respectively.

   a) Find the probability density function of $X + Y$.

   b) Find the conditional probability density function of X, given that $X + Y = m$.

4) Suppose that Y has a gamma distribution with $\alpha = \frac{n}{2}$ for some positive integer n and β equal to some specified value. Use the method of moment-generating functions to show that $W = \frac{2Y}{\beta}$ has a χ^2 distribution with n degrees of freedom.

5) Suppose that X and Y have gamma distributions with different α's but the same β, and that X and Y are independent. Let $U = \frac{X}{X+Y}$ and $V = X + Y$.

   a) Derive the joint density function for U and V.

   b) Show that the marginal distribution of U is a beta distribution with parameters α_x and α_y.

   c) Show that the marginal distribution of V is a gamma distribution with parameters $\alpha = \alpha_x + \alpha_y$ and β.

d) Establish that U and V are independent.

Homework 8, Due November 3

1) Let X and Y be independent and uniformly distributed over the interval $(0,1)$.

   a) Find the probability density function of $U = \min(X, Y)$.

   b) Find $E(U)$ and $V(U)$.

2) Let $Y_1, Y_2, \dots, Y_n$ be independent and uniformly distributed over the interval $[0, \theta]$.

   a) Find the probability distribution function of $U = \max(Y_1, Y_2, \dots, Y_n)$.

   b) Find the probability density function of U.

   c) Find the mean and variance of U.

3) The amount of fill dispensed by a bottling machine is normally distributed with $\sigma = 2$ ounces.

   a) If $n = 9$ bottles are randomly selected from the output of the machine, what is $P(|\bar{Y} - \mu| \leq .3)$?

   b) Find $P(|\bar{Y} - \mu| \leq .3)$ when $\bar{Y}$ is to be computed using samples of sizes $n = 25$, $n = 36, n = 49$, and $n = 64$.

   c) What pattern do you observe among the values for $P(|\bar{Y} - \mu| \leq .3)$ for the various values of n?

4) A forester studying the effects of fertilization on certain pine forests in the Southeast is interested in estimating the average basal area of pine trees. In studying basal areas of similar trees for many years, he has discovered that these measurements (in square inches) are normally distributed with standard deviation approximately 4 square inches. The forester would like the sample mean to be within 1 square inch of the population mean, with probability 0.90. How many trees must he measure in order to ensure this degree of accuracy?

5) The fracture strength of tempered glass averages 14 (measured in thousands of pound per square inch) and has standard deviation 2.

   a) What is the probability that the average fracture strength of 100 randomly selected pieces of this glass exceeds 14.5?

   b) Find an interval that includes, with probability 0.95, the average fracture strength of 100 randomly selected pieces of this glass.

Homework 9, Due November 10

1) A rough rule of thumb is that the normal approximation to the binomial distribution is adequate if $p \pm 3\sqrt{\frac{pq}{n}}$ lies in the interval $(0,1)$, that is, if $0 < p - 3\sqrt{\frac{pq}{n}}$ and $p + 3\sqrt{\frac{pq}{n}} < 1$.

   a) Show that $p + 3\sqrt{\frac{pq}{n}} < 1$ if and only if $n > 9\frac{p}{q}$.

   b) Show that $0 < p - 3\sqrt{\frac{pq}{n}}$ if and only if $n > 9\frac{q}{p}$.

   c) For what values of n will the normal approximation to the binomial distribution be adequate if $p = 0.5$?

   d) Repeat part c) for $p = 0.6, 0.4, 0.8, 0.2, 0.99,$ and 0.001.

2) Suppose that $E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$, $V(\hat{\theta}_1) = \sigma_1^2$, and $V(\hat{\theta}_2) = \sigma_2^2$. Consider the estimator $\hat{\theta}_3 = a\hat{\theta}_1 + (1-a)\hat{\theta}_2$.

   a) Show that $\hat{\theta}_3$ is an unbiased estimator for θ.

   b) If $\hat{\theta}_1$ and $\hat{\theta}_2$ are independent, how should the constant a be chosen in order to minimize the variance of $\hat{\theta}_3$?

3) Suppose $Y_1, Y_2,$ and Y_3 are a random sample from an exponential distribution with mean θ. Consider the following five estimators of θ:

   $\hat{\theta}_1 = Y_1, \quad \hat{\theta}_2 = \frac{Y_1+Y_2}{2}, \quad \hat{\theta}_3 = \frac{Y_1+2Y_2}{3}, \quad \hat{\theta}_4 = \min(Y_1, Y_2, Y_3), \quad \hat{\theta}_5 = \bar{Y}.$

   a) Which of these is an unbiased estimator for θ?

   b) Among the unbiased estimators, which has the smallest variance?

4) Let $Y_1, Y_2, \dots, Y_n$ be a random sample of size n from a population whose density is given by $f(y) = \begin{cases} \frac{\alpha y^{\alpha-1}}{\theta^\alpha} & 0 \le y \le \theta \\ 0 & \text{elsewhere} \end{cases}$, where $\alpha > 0$ is a known, fixed value, but θ is unknown. Consider the estimator $\hat{\theta} = \max(Y_1, Y_2, \dots, Y_n)$.

   a) Show that $\hat{\theta}$ is a biased estimator for θ?

   b) Find a multiple of $\hat{\theta}$ that is an unbiased estimator of θ.

   c) Derive $\text{MSE}(\hat{\theta})$.

Homework 10, Due November 21

1) Let $Y_1, Y_2, \dots, Y_n$ be a random sample of size n from a uniform distribution on the interval $[0, \theta]$. $\hat{\theta}_1 = (\mathrm{n}+1)\min(Y_1, Y_2, \dots, Y_n)$ and $\hat{\theta}_2 = \frac{n+1}{n}\max(Y_1, Y_2, \dots, Y_n)$ are both unbiased estimators for θ. Show that the efficiency of $\hat{\theta}_1$ to $\hat{\theta}_2$ is $\frac{1}{n^2}$. Notice that this implies that $\hat{\theta}_2$ is a markedly superior estimator.

2) Suppose $Y_1, Y_2, \dots, Y_n$ are a random sample of size n from a Poisson distribution with mean λ. Consider $\hat{\lambda}_1 = \frac{Y_1+Y_2}{2}$ and $\hat{\lambda}_2 = \bar{Y}$. Derive the efficiency of $\hat{\lambda}_1$ relative to $\hat{\lambda}_2$.

3) Let $Y_1, Y_2, \dots, Y_n$ be a random sample of size n from a standard normal distribution.

   a) What is the distribution of $\sum_{i=1}^{n} Y_i^2$?

   b) Let $W_n = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$. Does W_n converge in probability to some constant? If so, what is the value of the constant?

4) Let $Y_1, Y_2, \dots, Y_n$ be a random sample of size n from a Pareto distribution, whose density is given by $f(y) = \begin{cases} \alpha\beta^\alpha y^{-(\alpha+1)} & \beta \le y \\ 0 & \text{elsewhere} \end{cases}$, where $\beta > 0$ is a known, fixed value, but α is unknown.

   a) Show that $\prod_{i=1}^{n} Y_i$ is sufficient for α.

   b) A probability is a member of the exponential family if

   $$f(y|\theta) = \begin{cases} a(\theta)b(y)e^{-[c(\theta)d(y)]} & a \le y \le b \\ 0 & \text{elsewhere} \end{cases},$$

   where a and b do not depend on θ. Show that the Pareto distribution is in the exponential family.

   c) For a distribution in the exponential family, $\sum_{i=1}^{n} d(Y_i)$ is sufficient for θ. Does that statistic for the Pareto agree with your answer to part a)?

Homework 11, Due December 3

1) Let $Y_1, Y_2, \dots, Y_n$ be a random sample of size n from a distribution whose density is given by $f(y) = \begin{cases} \theta y^{\theta-1} & 0 \le y \le 1 \\ 0 & \text{elsewhere} \end{cases}$, where $\theta > 0$ is unknown.

   a) Show that this density is in the exponential family and that $\sum_{i=1}^{n} -\ln(Y_i)$ is sufficient for θ.

   b) Show that $W_i = -\ln(Y_i)$ has an exponential distribution with mean $\frac{1}{\theta}$.

c) Show that $2\theta \sum_{i=1}^{n} W_i$ has a χ^2 distribution with $2n$ degrees of freedom.

d) Show that $E\left(\frac{1}{2\theta \sum_{i=1}^{n} W_i}\right) = \frac{1}{2(n-1)}$.

e) What is the MVUE for θ?

2) Let $Y_1, Y_2, \dots, Y_n$ be a random sample of size n from a distribution whose density is given by $f(y) = \begin{cases} \frac{2}{\theta^2}(\theta - y) & 0 \le y \le \theta \\ 0 & \text{elsewhere} \end{cases}$.

a) Find an estimator for θ by using the method of moments.

b) Is this estimator a sufficient statistic for θ?

3) Let $Y_1, Y_2, \dots, Y_n$ be a random sample of size n from a distribution whose density is given by $f(y) = \begin{cases} (\theta + 1)y^{\theta} & 0 \le y \le 1, \theta > -1 \\ 0 & \text{elsewhere} \end{cases}$. Find the MLE for θ.

4) Let Y_1 and Y_2 be a random sample of size 2 from the uniform distribution over the interval $[\theta, \theta + 1]$. For testing $H_0\colon \theta = 0$ versus $H_a\colon \theta > 0$, we have two competing tests:

Test 1: Reject H_0 if $Y_1 > 0.95$.

Test 2: Reject H_0 if $Y_1 + Y_2 > c$.

Find the value of c so that test 2 has the same value for α as test 1.

5) The Rockwell hardness index for steel is determined by pressing a diamond point into the steel and measuring the depth of penetration. For 50 specimens of an alloy of steel, the Rockwell hardness index averaged 62 with standard deviation 8. The manufacturer claims that this alloy has an average hardness index of at least 64. Is there sufficient evidence to refute the manufacturer's claim at the 1% significance level?

Homework 12, Due December 10

1) Refer to 5) from Homework 11. The steel is sufficiently hard to meet usage requirements if the mean Rockwell hardness measure does not drop below 60. Using the rejection region found in 5) from Homework 11, find β for the specific alternative $\mu_a = 60$.

2) Refer to 4) from Homework 11.

a) Find the power of test 1 for the following θ's: .1, .4, .7, 1.0. Sketch a graph of the power function.

b) Find the power of test 2 for the following θ's: .1, .4, .7, 1.0. Sketch a graph of the power function.

c) Compare the two power curves. What can you conclude about the power of test 2 compared to the power of test 1 for *all* $\theta \geq 0$?

3) Let Y_1 be a random sample of size 1 from a distribution whose density is given by $f(y) = \begin{cases} \theta y^{\theta-1} & 0 \leq y \leq 1, \theta > 0 \\ 0 & \text{elsewhere} \end{cases}$.

   a) Sketch the power function of the test with rejection region: $Y > .5$.

   b) Based on the single observation Y, find a uniformly most powerful test of size α for testing $H_0: \theta = 1$ versus $H_a: \theta > 1$.

4) Let $Y_1, Y_2, \dots, Y_n$ be a random sample of size n from a Bernoulli distribution, whose density is given by $p(y_i) = p^{y_i}(1-p)^{1-y_i}, y_i = 0, 1$. Suppose we are interested in testing $H_0: p = p_0$ versus $H_a: p = p_a$, where $p_0 < p_a$.

   a) Find the most powerful test, and give the form of the rejection region.

   b) Recall that $\sum_{i=1}^{n} Y_i$ has a binomial distribution with parameters n and p. Indicate how to determine the values of any constants contained in the rejection region derived in part a).

   c) Is the test from part a) uniformly most powerful? Why or why not?