MTH U481 : SPRING 2009: PRACTICE PROBLEMS FOR FINAL

1). Two urns are provided as follows: urn 1 contains 2 white chips and 4 red chips, while urn 2 contains 5 white chips and 3 red chips. One chip is chosen at random, based on the toss of a biased coin which has probability 0.75 of heads. If the coin is heads, choose a chip at random from urn 1; if the coin is tails, choose a chip at random from urn 2. For convenience, W means white, R means red, H means heads and T means tails.

a). Draw a tree diagram to represent the sample space.

b). Find the four conditional probabilities $P(W|H)$, $P(W|T)$, $P(R|H)$, $P(R|T)$.

c). Find the probability that a white chip was picked.

d). Given that a white chip was picked, what is the probability that the coin was heads?

2). Two fair dice are thrown at random. Define the discrete random variable Y as follows: if the numbers on the dice are equal, then $Y = 1$; if the numbers on the dice are different, then $Y = 0$. Find the pdf of Y.

3). 20 bits are transmitted through a binary communication channel. Suppose that every bit has probability $p = 0.2$ to be changed as it crosses the channel – that is, a 0 may be changed to a 1 with probability 0.2, and a 1 may be changed to a 0 with probability 0.2. Find the probability that there are at least three changes in the transmission.

4). It is midnight on December 31, 1999. The lights go out in your house. However your flashlight works (it uses one battery), and you see that you have also three spare batteries for it. The lifetimes of these four batteries are random and independent, but you know that they all have the same mean (8 hours), and the same standard deviation (1/2 hour). Assume that when one of them fails you immediately replace it with a fresh battery. Let T be the total length of time that the flashlight will work. What is the mean and the standard deviation of T? [Hint: write T in terms of the lifetimes of the batteries]

5). In a certain neighborhood of Boston, it is observed that 1 in 10,000 houses has a serious fire every year. If there are 20,000 houses in the neighborhood, use the Poisson distribution to estimate the probability that two or more houses have serious fires in a given year.

6). A certain computer chip is known to have a lifetime which is a random variable; it has a normal distribution with mean 1000 hours and standard deviation 80 hours.

a) Find the probability that a randomly selected chip fails in less than 900 hours (use the table provided for the standard normal distribution).

b) A second chip has just come on the market, whose lifetime is also a normal random variable, with mean 1050 hours and standard deviation 150 hours. Find the probability that this new chip fails in less than 900 hours.

c) You decide to use a mixture of the two types of chip in your next computer – you use 40% of the first type, and 60% of the second. If you pick a chip at random from this mixture, what is the probability that the chip fails in less than 900 hours? [Hint: use the results from (a) and (b), and your knowledge of conditional probabilities.]

7). A manufacturing process produces widgets whose length is a random variable, with pdf

$$f(x) = \begin{cases} kx^2 & \text{for } 0 \le x \le 1 \\ k & \text{for } 1 \le x \le 2 \end{cases}$$

where k is a numerical constant. Let $\overline{X}$ be the average length of a batch of 200 widgets.

a). Find the value of k.

b). Find the mean and standard deviation of $\overline{X}$.

c). Use the Central Limit Theorem to estimate the probability that $\overline{X}$ exceeds 1.34.

8). Five resistors are put in series in a circuit, so the total resistance is the sum of their resistances. The resistances are all normal random variables, with mean 100 Ohms and standard deviation 5 Ohms. Assuming the resistors are independent, what is the probability that the total resistance is more than 520 Ohms?

9). The two independent random variables X and Y are both uniformly distributed on the interval $[0, 1]$.

a). Sketch the sample space for the pair (X, Y), that is the region in the xy-plane where the joint pdf is non-zero.

b). Calculate $P(|X - Y| \leq 1/2)$.

c). Find $\mathrm{E}[X - 2Y]$ and $\mathrm{VAR}[X - 2Y]$.

d). A large asteroid called Doom is orbiting around the Sun, and it may hit the Earth on April 15 in the year 2108. Both the Earth and Doom will arrive in the danger zone between 11:00 am and 11:20 am. If they arrive less than ten minutes apart, a collision will occur (technically speaking, Earth will be "toast"). Assuming that both arrival times are independent and uniformly distributed between 11:00 am and 11:20 am, find the probability that a collision will occur.

10). A certain brand of car has a defect which may cause it to break down after 30,000 miles. It is known that this problem occurs for one out of every eight of these cars. A rental company owns 250 of these cars. Use the normal appoximation to estimate the probability that more than 40 of their cars will break down because of this defect.

11). The continuous random variable X has a pdf depending on one parameter θ. Suppose that $\widehat{\theta_1}$ and $\widehat{\theta_2}$ are independent estimators of θ, and both are unbiased. Suppose also that the mean square errors are

$$\mathrm{MSE}[\widehat{\theta_1}] = 12, \quad \mathrm{MSE}[\widehat{\theta_2}] = 20$$

Find the mean square error of the estimator

$$\widehat{\theta_3} = a\widehat{\theta_1} + (1 - a)\widehat{\theta_2}$$

and find the value of a which minimizes it.

12). In a recent poll of 2000 Americans, 75% expressed their belief that global warming is a very important problem facing the world. Construct a 99% confidence interval for the fraction of the American population that is concerned about global warming.

13). The capacities (in ampere-hours) of 10 batteries were recorded as follows:

$$140, 136, 150, 144, 148, 152, 138, 141, 143, 151$$

Find a value v that enables us to state with 90% confidence that the mean capacity is less than v (assume the capacities follow a normal distribution). [Helpful facts: let y_i be the recorded capacities, then $\sum y_i = 1443$ and $\sum y_i^2 = 208515$].

14). Suppose the null hypothesis $H_0 : \mu = 120$ is tested against the alternative $H_1 : \mu \neq 120$ for 16 samples from a normal distribution whose standard deviation is $\sigma = 10$. Find the p-value associated with the sample mean $\overline{y} = 122.3$.

15). $X_1, X_2, \ldots, X_n$ are independent random variables each with pdf

$$f_X(x) = \theta\, x^{\theta-1}, \quad 0 < x < 1$$

a) Find the maximum likelihood *estimator* for θ.
b) Find the maximum likelihood *estimate* if we observe $X_1 = 0.2$, $X_2 = 0.5$ and $X_3 = 0.4$.

16). The independent sample $X_1, \ldots, X_n$ is drawn from a distribution which depends on an unknown parameter θ. The mean and variance of this distribution are

$$\mathrm{E}[X] = \theta, \quad \mathrm{VAR}[X] = \theta^2 + 1$$

The estimator for θ is $\widehat{\theta} = c\overline{X}$, where $\overline{X} = 1/n\ (X_1 + \cdots + X_n)$ is the sample mean, and c is a number which you will determine.

a). Calculate the bias of the estimator $\widehat{\theta}$ as a function of c.

b). Find $\mathrm{MSE}[\widehat{\theta}]$, the mean square error of $\widehat{\theta}$ (your answer will depend on n, c and θ).

c). Find the value of c which minimizes $\mathrm{MSE}[\widehat{\theta}]$ (your answer will depend on n and θ). What happens to this value as $n \to \infty$?

17). From past experience, the standard deviation for the number of beans in a one pound bag of Starbucks coffee is 100 beans.
a) A quality control officer tests 50 one pound bags and the average number of beans is 1074 in the tested bags. Find a 95% confidence interval for the true average number of beans in all one pound bags.
b) If the officer wishes to estimate the true average number of beans in all one pound bags to an accuracy of 10 beans at 96% confidence level, how many bags should be tested?
c) Starbucks Coffee Roasters try to get an average of 1100 beans in a pound of coffee. If the true average is 1070, what is the probability that in a test of 50 bags, the 99% confidence interval would not contain 1100?

18). Wonder Drug Company claims that its latest product has dangerous side effects in "only" 40% of all patients. In a clinical trial with 300 patients taking the product, 130 patients have side effects. Find a 94% two-sided confidence interval for p, the *proportion among all patients* who experience side effects from the drug. Also find the p-value associated with this test. Is the Companys claim consistent with the clinical trial data?

19). Peter tosses a biased coin repeatedly until a head comes up. Let p denote the probability that a head comes up in one toss.
a) Find the pdf for X, the number of tosses until the first head appears.
b) If the first head comes up on the $X = 4$ toss, find the MLE for p.

20). The mean length of pregnancy for a house cat is 60 days. A sample of 10 feral cats finds a mean length of 53 days with a standard deviation of 7. Test at the 5% level of confidence to see if the mean length for feral cats is less.
a) Give the null and alternate hypotheses.
b) Find the test statistic, decide if you accept or reject the null hypothesis, and explain what that means here.

21). Echinacea was tested as a treatment for upper respiratory infections. A sample of 337 children given echinacea had fever for a mean of 0.81 days with a standard deviation of 1.5. A sample of 370 children given a placebo had fever for a mean of 0.64 days with a standard deviation of 1.16. Test at the 5% level to see if echinacea changed the length of time with fever.
a) Give the null and alternate hypotheses.
b) Find the p-value, decide if you accept or reject the null hypothesis, and explain what that means here.

22). The ages of a random sample of seven Mathematics Professors at WestSouthern University are 39, 54, 61, 72, 59, 44, 35. Find a 90% confidence interval for the standard deviation of the ages of all Mathematics Professors at WSU.

23). A company wants to see if an ad campaign will increase sales of a product. They will assume standard deviation $\sigma = 100$ per day, and the null assumption is H_0: $\mu = 400$. The decision rule is that they will reject H_0 if a sample of 50 days finds an average of more than 425 (so reject H_0 if $\bar{X} > 425$). Find P(reject H_0 if H_0 is true).