Probability problems.

1. A jar has 10 black and 5 white balls. Four balls are taken out one by one at random, without replacement. Find the probability that the first 2 balls had different colors, given that 2 white and 2 black balls were taken out.

3. You sent an invitation to a party to 3 of your friends by email. Your friends reply independently. The waiting times for your friends' replies are exponential with parameters 1, 2, and 3, respectively.

(a) Find the distribution of the time you need to wait until the first reply.

(b) Find the probability that you receive a reply from the second friend before the reply from the third friend.

4. A bakery makes X cakes, where X is geometric with parameter $0 < p < 1$. All cakes are sold today with probability p, all cakes but one are sold today with probability $1-p$. Let Y be the number of cakes sold today.

(a) Find $E(Y)$.

(b) Find the probability mass function of Y.

6. Let X and Y be independent Poisson random variables with parameters λ and 2λ, respectively. Find the conditional distribution of X given $X+Y$.

Statistics Problems: Put all answers on these sheets

1) Company XXX has 200 employees. In order to estimate the number of automobiles owned by these employees the company takes a simple random sample of 80 of them, and determines that 22 have no auto, 40 have one auto, 14 have two autos, and 4 have 3 autos.

a) What is meant by "simple random sample"?

b) Find the sample mean $\bar{X}$ and sample variance S^2 for the numbers of autos owned by the 80 employees sampled.

c) Estimate $Var(\bar{X})$.

d) Find an approximate 95% confidence interval on the population total T of all autos owned by the 200 employees.

e) Explain what is meant by "95% confidence interval" so your friend would understand. Your friend understands the meanings of "population and sample means", and of "sample variances," but has never heard of "confidence intervals."

2. Let $(X_1, X_2, \ldots, X_{25})$ be the values observed when a random sample of 25 is taken from $N(\mu, \sigma^2)$, with $\sigma^2 = 100$ known.

(a) Give the uniformly most powerful $\alpha = .05$ test of H_0: $\mu = 40$ versus H_1: $\mu > 40$. Just give the test. You need not prove that it is uniformly most powerful.

(b) Give the power of the test for $\mu = 45$.

(c) What must n be so that the power at $\mu = 42$ is .80?

(d) Suppose that σ was **not** known, that $\overline{X}$ was observed to be 43.0 and S^2 (the sample variance) was 81.0. Find a 99% confidence interval on μ.

3. Let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution with density

$$f(x\,;\theta) = \theta x^{-\theta-1} \text{ for } x \geq 1, \theta > 1$$

a) Find the method of moments estimator $\hat{\theta}_n$ of θ.

b) What does it mean to say that a sequence $\{\hat{\theta}_n\}$ of estimator of a parameter θ is a **consistent sequence** of estimators of θ?

c) Is the sequence $\{\hat{\theta}_n\}$ you obtained in a) consistent for θ? Why?

d) Find the maximum likelihood estimator of θ.

4. Suppose that X_1 and X_2 are independent, with the same probability mass function f_0 or f_1, where f_0 and f_1 are as follows:

k:	0	1	2
$f_1(k)$	0.6	0.3	0.1
$f_0(k)$	0.1	0.2	0.7

a) Use the table below to determine the likelihood ratio $\Lambda(x_1, x_2)$ for all possible samples. You can use the cells of the table to do the necessary computations.

			x_2	
		0	1	2
	0			
x_1	1			
	2			

(b) Of all possible tests of size $\alpha = 0.09$, which is most powerful?

(c) What is the power of the most powerful test of size $\alpha = 0.09$?

(d) Give a test of the same size or greater size but with less power than that of the Neyman-Pearson test.

5. Consider the standard normal theory regression model with regression function $E(Y \mid x) = \beta_0 + \beta_1 x + \beta_2 x^2$ and the data

x	-1	-1	0	1	1
y	1	3	3	7	5

(a) Find the least squares estimates of the parameters β_0, β_1, *and* β_2.. (If you know how to invert a 2 by 2 matrix, you should be able to invert the 3 by 3 matrix to be used in this case.)

(b) Determine SSE.

(c) Estimated the variance of $\hat{\beta}_2$.

(d) Give a 95% CI estimate on β_2.

6. A crop scientist had 5 plots of land available for experimentation. In order to determine which of two seed, A or B, would produce the greatest yield of corn, he divided each plot into two parts, and randomly assigned the A to one part, B to the other. The plots and yields in kilograms of corn were:

Plot:	1	2	3	4	5
A	9	12	6	7	8
B	7	9	4	7	5

The scientist wishes to decide whether the new seed A is better than the old seed B. His plan is to continue to use B unless there is strong evidence that A is better. State a model, state null and alternative hypotheses, and then carry out an appropriate one-sided test. Use $\alpha = 0.05$.

7. A rich man decides to divide his fortune, which consists of N distinct gold coins, among his m friends, $N > m \geq 1$.

   (a) In how many distinct ways can the coins be distributed to his friends?

   (b) In how many ways can the coins be distributed if every friend must receive at least one?

8. A hen lays X eggs, where X is Poisson with parameter λ. Each hatches with probability p, independently of the others, yielding Y chicks.

   (a). Show that Y is a Poisson distribution with parameter λp.

   (b). Show that $\rho(X, Y) = \sqrt{p}$.

9. Let $X_1, X_2, \ldots$ be independent, identically distributed (discrete) random variables with an unknown mass function $f(x)$. For a fixed x, how can we use the weak law of large numbers to approximate $f(x)$?

10. Expensive patented manufactures are often copied, and the copies sold as genuine. You are replacing a part in your car; with probability p you buy a false part, which probability $1-p$ a genuine part. In each case lifetimes are exponential, false parts with parameter μ, genuine parts with parameter λ, where $\lambda < \mu$. The life of the part you install is T. Compute the hazard rate function $r(t) = f(t)/(1 - F(t))$ of T. Does it make a difference if $\lambda > \mu$?

11. Let X and Y have the uniform density over the unit circular disk C, namely

$$f(x, y) = \frac{1}{\pi}, \quad (x, y) \in C.$$

   (a) Are X and Y independent?

   (b) Find $f_X(x)$ and $f_Y(y)$.

   (c) If $X = R\cos\theta$, and $Y = R\sin\theta$, are R and θ independent?

Statistics Problems

12. (5 points each part) Let $X_1, \ldots, X_n$ be a random sample from a distribution F with pdf

$$f(x, \theta) = \frac{2x}{\theta} e^{-x^2/\theta}, \quad x > 0.$$

   (a). Show that X_1^2 is exponential with density function

$$f_{X^2}(x, \theta) = \frac{1}{\theta} e^{-x/\theta}, \quad x > 0.$$

   (b). Show that the distribution of $\sum_{i=1}^n X_i^2$ is Gamma$(n, 1/\theta)$, i.e., it has the density function

$$G(x; n, \theta) = \frac{1}{\theta^n \Gamma(n)} x^{n-1} e^{-x/\theta} \quad \text{for } x > 0.$$

   [Hint: you may use the moment generating function argument. The moment generating function of Gamma$(n, 1/\theta)$ is $(1 - \theta t)^{-n}$.]

   (c). Find the method of moment estimator $\widehat{\theta}_1$ of θ.

(d). Find the maximum likelihood estimator $\widehat{\theta}_2$ of θ. Show that $\widehat{\theta}_2$ is unbiased for θ.

(e). Find a sufficient statistic for θ.

(f). Find a $100(1-\alpha)\%$ confidence interval for θ using χ^2 percentiles.

(g). Derive a uniformly most power test of size α for the test of H_0: $\theta \leq \theta_0$ against H_a: $\theta > \theta_0$

13. (10 points) Let $X_1, \ldots, X_n$ be n independent observations and $X_i \sim \text{Poisson}(\lambda_i)$. Derive via the generalized likelihood ratio a test of size α for

$$H_0 : \text{all } \lambda_i \text{ are equal} \qquad \textit{versus} \qquad H_a : \text{ not all equal.}$$

Provide the rejection region.

14. (15 points) A study is done to compare the performances of engine bearings made of different compounds. Ten bearings of each type were tested. The following are the times until failure:

Type I	3.03	5.53	5.60	9.30	9.92	12.51	12.95	15.21	16.04	16.84
Type II	3.19	4.26	4.47	4.53	4.67	4.69	12.78	6.79	9.37	12.75

$$\overline{x} = 10.69, \quad s_x^2 = 23.23, \quad \overline{y} = 6.75, \quad s_y^2 = 12.98$$

(a). Use normal theory to test H_0: there is no difference between the two types of bearings.

(b). Test the same H_0 using a nonparametric method.

15. (15 points) Consider the linear model $Y_{n\times 1} = X_{n\times p}\beta_{p\times 1} + \epsilon_{n\times 1}$. Assume that the columns of X are independent and $E(\epsilon_i) = 0$ $(i = 1, 2, \ldots, n)$.

(a). Find the LS estimator $\widehat{\beta}$ of β.

(b). Show that $\widehat{\beta}$ is unbiased.

16. (15 points) Suppose that you wished to estimate the proportion p_A of $N = 10000$ voters who favor candidate A in an election with one other candidate B.

(a). If sampling were with replacement, how large must the sample size n be in order to estimate p_A within .02 with probability at least .95?

(b). How large must n be if sampling is without replacement?

(c). If $n = 1000$ and 553 favor A, give a 95% confidence interval on the difference $\Delta = p_A - p_B$.