STAT 3010 Fall 2010 Exam #1

Instructor: Erkan Nane

Name:
PID:

Instructions

1. Read each problem carefully. Show all your work. Credit will only be awarded if your work is included.
2. There is a formula sheet at the end!!.

Problem	max-points	Points
1	20	
2	20	
3	30	
4	30	
Total	100	

Problem 1. x =the actual tracking weight of a stereo cartridge set to track at 3 g can be regarded as a continuous variable with density function

$$f(x) = \begin{cases} c[1-(x-3)^2], & 2 \le x \le 4 \\ 0, & otherwise. \end{cases}$$

a. Determine the value of c [you might find it helpful to graph $f(x)$].

b. What proportion of actual tracking weights exceed the target weight?

c. What proportion of actual tracking weights are within 0.25 g of the target weight?

d. What is the mean and median values of x.

Problem 2. A plot to asses the plausibility of an exponential population distribution can be based on quantiles of the exponential distribution having $\lambda = 1$ (i.e., the exponential distribution with density function $f(x) = e^{-x}$ for $x > 0$). This is because λ, like σ for a normal distribution, is a scale parameter. Consider the following failure time observations (1000s of hours) resulting from accelerated life testing of 5 integrated circuit chips of a certain type:

$$242.0,\ 244.8,\ 270.0,\ 307.8,\ 359.5$$

Construct a quantile plot and comment on the plausibility of failure time having an exponential distribution.

Problem 3. The accompanying data on $x =$ current density (mA/cm^2) and $y =$ rate of deposition (μm) is as follows:

x	20	40	50	80
y	0.24	1.20	1.71	2.22

Summary quantities are $\sum x = 190, \sum y = 5.37, \sum x^2 = 10900, \sum y^2 = 9.35, \sum xy = 315.9$.

a. Construct a scatter plot of data.

b. Comment on form, strength and direction of the relationship between x and y.

c. Calculate the value of the (Pearson) correlation coefficient.

d. What percentage of the variation in y is explained by the approximate linear relationship between the two variables?

e. Determine the equation of the least squares line.

f. Predict a value for y (rate of deposition) using the least squares line for $x = 30$

Problem 4. A large insurance agency services a number of customers who have purchased both homeowner's policy and an automobile policy. For each type of policy, a deductible amount must be specified. Let x denote the homeowner's deductible amount and y denote the automobile deductible amount for a customer who has both type of policies. The joint mass function of x and y is as follows:

		y		
f(x,y)		0	300	400
	100	0.20	0.10	0.20
x	400	0.05	0.15	0.30

a. What proportion of customers have $400 deductible amounts for both types of policies?

b. What is the marginal mass function of x? What is the marginal mass function of y?

c. Are x and y independent?

d. Compute the covariance between x and y and then the value of the correlation coefficient. Do these two variables appear to be strongly correlated? Explain.

Formulae

$$\mu_x = \begin{cases} \int x f(x)dx, & x \text{ continuous} \\ \sum x p(x), & x \text{ discrete.} \end{cases}$$

$$\sigma_x^2 = \begin{cases} \int x^2 f(x)dx - (\mu_x)^2, & x \text{ continuous} \\ \sum x^2 p(x) - (\mu_x)^2, & x \text{ discrete.} \end{cases}$$

If x and y are discrete and they have a joint mass function $f(x, y)$ then for a function $h(x, y)$

$$\mu_{h(x,y)} = \sum_{x,y} h(x, y) f(x, y);$$

covariance of x and y is

$$cov(x, y) = \mu_{xy} - \mu_x \mu_y;$$

and population correlation coefficient is

$$\rho = \frac{cov(x, y)}{\sigma_x \sigma_y}.$$

$$\bar{x} = \frac{\sum x_i}{n}, s^2 = \frac{1}{n-1}\left(\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}\right) = \frac{S_{xx}}{n-1}$$

$$S_{xx} = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n},\ S_{yy} = \sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}$$

$$S_{xy} = \sum x_i y_i - \frac{\left(\sum x_i\right)\left(\sum y_i\right)}{n}$$

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

$$b = \frac{S_{xy}}{S_{xx}},\ a = \bar{y} - b\bar{x}$$