Exam Statistics, Probability and Calculus

January 15, 2010, 10:00-13:00 h.

Use of Rice's book, class notes, calculator and laptop is allowed. If you perform any computations with `R`, please provide the commands you used.

1. (a) Assume we have a joint probability mass function of X and Y

$$p_{X,Y}(x,y) = p^2(1-p)^{x+y-2}, \quad x = 1, 2, \dots \text{ and } y = 1, 2, \dots$$

and $p_{X,Y}(x,y) = 0$ otherwise, where $0 < p < 1$. Find the marginal distribution of X. Do you know the name of this distribution?

(b) Now consider two light bulbs A and B, with lifetimes X and Y. Suppose the joint density of X and Y is

$$f_{X,Y}(x,y) = \lambda\mu e^{-(\lambda x + \mu y)}, \quad x, y > 0$$

and $f_{X,Y}(x,y) = 0$ otherwise. Here λ and μ are positive constants. Determine the probability that both lamps are still burning at time t.

(c) Determine the probability that lamp A fails first.

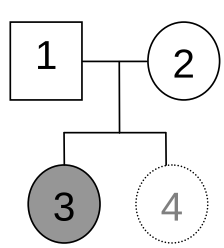
2. Consider the genotypes AA, AB and BB with frequencies 0, 0.02 and 0.98 respectively. Persons carrying genotype AB have probability 0.85 to have a disease D. Persons with genotype BB have probability of 0.03 to have disease D. Define the random variable X to be 1 if a person is AB and 0 otherwise and the random variable Y to be 1 if a person has disease D and to be 0 if a person has not the disease.

(a) Find the prevalence of disease D in the population.

(b) Find the variance of Y conditional on $X = 1$.

(c) Find the variance of Y.

(d) In the family below, numbers 1 and 2 are healthy and number 3 has disease D. Person 2 is pregnant of and wants to know the probability for her newborn 4 to have the disease.



Find the probability that number 4 has the disease. Hint: Start with the genotypes of the parents and notice that the probability of more than 1 copy of the D allele is very small.

(e) Consider a randomly chosen pair of two siblings. Let X_1 be the number of A alleles carried by sibling 1 and let X_2 be the number of A alleles carried by sibling 2. The joint distribution is given by

	$X_1 = 0$	$X_1 = 1$
$X_2 = 0$	0.97	0.01
$X_2 = 1$	0.01	0.01

Find the covariance of X_1 and X_2.

(f) Consider two quantitative variables $Y_1 = X_1 + U + V$ and $Y_2 = X_2 + U + W$ with U, V and W independent normally distributed random variables with variance equal to 1. We have

$$\mathrm{cov}(X_1, U) = \mathrm{cov}(X_1, V) = \mathrm{cov}(X_1, W) = \mathrm{cov}(X_2, U) = \mathrm{cov}(X_2, V) = \mathrm{cov}(X_2, W) = 0.$$

For the covariance between X_1 and X_2, see question (e). Now, find the covariance between Y_1 and Y_2.

3. Let $X_1, X_2, \ldots, X_n$ be independent random variables with density function

$$f_\theta(x) = (\theta + 1)x^\theta, \qquad 0 \le x \le 1$$

(a) Find the method of moments estimator of θ.

(b) Find the maximum likelihood estimator of θ.

(c) Find the asymptotic variance of the MLE.

4. Let $X_1, \ldots, X_{10}$ be a random sample of size 10 from a Bernoulli distribution, which is given by

$$P(X = 0) = p \quad \text{and} \quad P(X = 1) = 1 - p,$$

where $0 \le p \le 1$ is an unknown parameter. Suppose the realization of our sample is

0 1 1 0 1 0 1 1 1 0

(a) We want to test

$$H_0 : p = 1/3 \ \text{ versus } \ H_A : p = 1/2.$$

Determine the most powerful test with significance level $\alpha = 0.05$. Do you reject H_0 on the basis of our sample?

(b) To simplify the calculations of the previous problem, you could have used the normal approximation. How?

(c) Compute the p value.