Marco Longfils and Petter Mostad
Applied Mathematics and Statistics
Chalmers and GU

MVE055 / MSG810 Matematisk statistik och diskret matematik

Exam 24 October 2017, 8:30 - 12:30

Allowed aids: Chalmers-approved calculator
and one (two-sided) A4 sheet of paper with your own notes.
Total number of points: 30. To pass, at least 12 points are needed.
Note: All answers should be motivated.

1. **(6 points)** Assume a project group with 4 persons is being put together from the staff of a department, where 8 women and 5 men work. Assume persons are selected into the group in a way that is unrelated to their gender.

   (a) What is the probability that the group will contain exactly 2 men and 2 women?

   (b) What is the probability that it will contain no men?

   (c) Assume persons are asked to join the group sequentially. What is the probability of observing the following sequence of genders of those who are invited: Female, female, male, male.

2. **(5 points)** Assume $X_1, \ldots, X_n$ is a random sample from Normal(μ_x, σ_x), i.e. from a normal distribution with expectation μ_x and variance σ_x^2. Assume $Y_1, \ldots, Y_m$ is a random sample from Normal(μ_y, σ_y). Assume that σ_x and σ_y are known, i.e., not estimated from the X_i or Y_i, but known from other sources. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i$.

   (a) Write down the formula for the probability distribution of $\overline{X}$.

   (b) Write down the formula for the probability distribution of $\overline{X} - \overline{Y}$.

   (c) Write down a random variable on the form

   $$\frac{\overline{X} - \overline{Y} - (\mu_x - \mu_y)}{\ldots}$$

   so that it has a standard normal distribution, i.e. a normal distribution with expectation 0 and variance 1.

   (d) Find formulas for random variables L_1 and L_2 so that $[L_1, L_2]$ becomes a 95% confidence interval for $\mu_x - \mu_y$.

3. **(5 points)** The World Health Organization has estimated that a certain disease is present in 1% of the population of a country. Doctors have the possibility to perform a test to check if a person has the disease. However, the test is not infallible. If a person has a disease, then the test results positive in 80% of the cases. If the person is healthy, then the test is positive in 5% of the cases.

(a) If we perform the test on a randomly selected person, what is the probability that the test will be positive?

(b) What is the probability that a person has the disease if the test resulted positive?

4. **(5 points)** Let $X_n, n = 0, 1, \dots,$ count the number of individual at time n in a population. Assume that $X_0 = 2$ and that the resources available in the environment only allow for a maximum number N of individuals. Assume that at each time n if $0 < X_n < N$ there could be either a birth, with probability $p = 0.25$, or a death, with probability 0.75. If the maximum number of individuals has been reached at time n, that is $X_n = N$, then it is only possible to have either a death, again with probability 0.75, or nothing happens (neither deaths or births). On the other hand, if the population goes extinct, that is $X_n = 0$, then there is no possibility for births and deaths.

   (a) Find the transition matrix P of the Markov chain $\{X_n\}_{n\in\mathbb{N}}$;

   (b) For $N = 2$, compute the expected number of steps before extinction.

   (c) Consider $Y_n = \max\{X_0, ..., X_n\}$. Is $\{Y_n\}_{n\in\mathbb{N}}$ a Markov chain? Motivate.

5. **(4 points)** Suppose that a company receives orders from two clients, A and B, independently of each other. Let X_A and X_B denote the random variables that count the number of orders received respectively from A and B in a week. Assume that X_A follows a Poisson distribution with parameter λ_A and X_B follows a Poisson distribution with parameter λ_B.

   (a) Let p be the probability that an orders has a problem and assume that each order has or has not a problem independently of the other orders. Let Y denote the number of orders with a problem. What is the probability that y orders have a problem in a week in which we have received n orders?

   (b) Denote by X the random variable which counts the total number of orders received in a week. Show that the density function of X is

$$P(X = n) = e^{-(\lambda_A+\lambda_B)}\frac{(\lambda_A + \lambda_B)^n}{n!}.$$

   Hint: it could be useful the following formula

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{(n-k)} \tag{1}$$

   (c) If in a week the company has receveid n orders, what is the probability that k of them has been done by A? That is, compute

$$P[X_A = k|X = n].$$

6. **(5 points)** Assume the continuous random variable has a probability distribution with expectation μ and variance σ^2, and assume $X_1, \dots, X_n$ is a random sample from this distribution.

(a) Write down the formula for the standard estimator for σ^2 in terms of $X_1, \ldots, X_n$.

(b) Compute in terms of μ and σ^2 the expectation

$$E\left[(X_i - \overline{X})^2\right]$$

where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, for any $i = 1, \ldots, n$.

(c) Prove that the estimator you presented in (a) is unbiased.