THE UNIVERSITY OF NEW SOUTH WALES

DEPARTMENT OF STATISTICS

Additional exercises for MATH5905, Statistical Inference

Part four: Multinomial distribution. Order statistics. Robustness

1. a) X_1, X_2, X_3 has a multinomial $(8; 0.2, 0.3, 0.5)$ distribution. Find $P(X_1 = 2, X_2 = 2, X_3 = 4)$, $E(X_2), Var(X_2)$ and $Cov(X_1, X_3)$.

   b) X_1, X_2, X_3 has a multinomial $(6; 0.5, 0.2, 0.3)$ distribution. Find $P(X_1 = 3, X_2 = 1, X_3 = 2)$ and $P(X_1 + X_2 = 2)$.

2. Find the probability density function of the second order statistic $X_{(2)}$ in a random sample of size four from a population with the density function

$$f(x) = \begin{cases} e^{1-x}, 1 < x < \infty, \\ 0 \text{ elsewhere} \end{cases}$$

3. Find the probability density function of $X_{(4)}$ in a random sample of size five from a population with the density function

$$f(x) = \begin{cases} \frac{1}{x^2}, 1 \leq x < \infty, \\ 0 \text{ elsewhere} \end{cases}$$

4. The opening prices per share of two similar stocks Y_1 and Y_2 are independent random variables, each with density function

$$f(y) = \begin{cases} \frac{1}{2}e^{-\frac{y-4}{2}}, y \geq 4, \\ 0 \text{ elsewhere} \end{cases}$$

   On a given morning Mr. A is going to buy shares of whichever stock is less expensive. Find the probability density function and the expected value for the price per share that Mr. A will have to pay.

5. Find the expected value of the largest order statistic in a random sample of size 3 from:

   a) the exponential distribution with density $f(x) = e^{-x}, x \geq 0$.

   b) the standard normal distribution

6. Electric components of a certain type have lifetime Y with probability density given by

$$f(y) = \begin{cases} \frac{1}{100}e^{-\frac{y}{100}}, y > 0, \\ 0 \text{ elsewhere} \end{cases}$$

   a) Suppose that two such components operate independently and in series in a certain system (that is, the system fails when either component fails). Find the density function for X, the lifetime of the system.

   b) Now suppose that the components operate in parallel (that is, the system does not fail until both components fail). Find the density function for X, the lifetime of the system.

7. A continuous random variable X has a standard exponential distribution

$$f(x) = \begin{cases} e^{-x}, x > 0, \\ 0 \text{ elsewhere} \end{cases}$$

   For a random sample of size 3, let $X_{(1)}, X_{(2)}, X_{(3)}$ denote the ordered sample.

a) Write down the joint distribution of $X_{(1)}$ and $X_{(3)}$.
b) Obtain the distributions of $X_{(1)}$ and $X_{(3)}$.
c) Evaluate $EX_{(1)}$ and $EX_{(3)}$.
d) Find the sampling distribution of the range $R = X_{(3)} - X_{(1)}$.

8. A random sample of size 3 is taken from a population with density

$$f(x) = \begin{cases} 2x, 0 \leq x < 1, \\ 0 \text{ elsewhere} \end{cases}$$

Find the sampling distribution of the range R.

9. For a random sample of size 2 from a standard normal distribution, find the distribution of the range.

10. The Cauchy-Schwartz Inequality tells us that for random variables X, Y with $E(X^2) < \infty$ and $E(Y^2) < \infty$, $\{E(|XY|)\}^2 \leq E(X^2)E(Y^2)$ holds. Using this Inequality, show that when estimating the location parameter θ in $f(x-\theta)$ by using an M-estimator (defined by its ψ function):

a) we get a loss in asymptotic efficiency in comparison to the MLE estimator (i.e. the Inequality

$$\sigma^2(F, \psi) = \frac{\int \psi^2(x) f(x) dx}{(\int \psi'(x) f(x) dx)^2} \geq \frac{1}{\int [\frac{f'(x)}{f(x)}]^2 f(x) dx}$$

holds).

b) Equality holds only when the M-estimator coincides with the MLE. (**But: despite the above observation we still may decide to use the M-estimator instead of the MLE because of the better robustness properties of the former in comparison to the latter**).

11. Assume that we are estimating the location parameter θ_0 for the family $f_\theta(x) = f(x-\theta)$ in a robustness context. The "true" value θ_0 to be estimated in robustness context is usually taken as the solution of the Equation $E_{\theta_0}\psi(X-\theta_0) = 0$. Derive the asymptotic normality statement from the lecture notes for the M-estimator a defined as the solution to the equation $\sum_{i=1}^n \psi(X_i - \hat{\theta}_M) = 0:$

$$\sqrt{n}(\hat{\theta}_M - \theta_0) \to^d N(0, \frac{\int \psi^2(x) f(x) dx}{(\int \psi'(x) f(x) dx)^2}).$$

Answers

1 a) 0.0945, 2.4, 1.68, -0.8 b) 0.135 0.059535

2) $12exp(3(1-x_{(2)}))(1-exp(1-x_{(2)})), 1 \leq x_{(2)} < \infty$

3) $20(x_{(4)}-1)^3/x_{(4)}^6, 1 \leq x_{(4)} < \infty$

4) $exp(-y_{(1)}-4), y_{(1)} \geq 4$. The expected value is 5.

5) a) 11/6, b) $\frac{3}{2\sqrt{\pi}}$.

6) a) $\frac{1}{50}exp(-\frac{x}{50})$; b) $\frac{1}{50}(1-exp(-\frac{x}{100}))exp(-\frac{x}{100})$

7) a) $6(e^{-x_{(1)}} - e^{-x_{(3)}})e^{-(x_{(1)}+x_{(3)})}, 0 < x_{(1)} < x_{(3)} < \infty$.

b) $3exp(-3x_{(1)}), x_{(1)} > 0; 3(1-exp(-x_{(3)}))^2 exp(-x_{(3)}), x_{(3)} > 0$.

c) $1/3, 11/6$; d) $2exp(-u)(1-exp(-u)), u > 0$.

8) $12u(1-u)^2, 0 < u < 1$.

9) $\frac{1}{\sqrt{\pi}}exp(-\frac{u^2}{4}), u > 0$.