

HG  
Nov. 08

## ECON 4130 H08

### Extra exercises for no-seminar week 47

(Solutions will be put on the net at the end of the week)

#### Exercise

(A slightly altered version of *Exam 2005H* -“*utsatt prove*”.)

#### Problem 1

- a. Let  $X$  and  $Y$  be two continuous random variables, both varying in the interval  $[0, 1]$ . The joint cumulative probability function (cdf) of  $X$  and  $Y$  is given by

$$F(x, y) = xy[1 + (0.8)(1 - x)(1 - y)] \quad \text{for } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1$$

Show that the joint density function for  $X$  and  $Y$  is given by

$$f(x, y) = 1 + (0.8)(1 - 2x)(1 - 2y) \quad \text{for } 0 \leq x \leq 1 \text{ og } 0 \leq y \leq 1.$$

- b. Show that both  $X$  and  $Y$  marginally are uniformly distributed over  $[0, 1]$ . Find the expectation and variance for  $X$  and  $Y$ . Are  $X$  and  $Y$  stochastically independent? Calculate  $P(Y \leq 0.5)$  and  $P(Y \leq 0.5 | X \leq 0.5)$ .

- c. Sketch the conditional density for  $Y$ ,  $f(y | x)$ , in a graph when  $x = 1/2$ , and  $x = 1$  respectively (i.e. altogether two graphs).

Calculate the regression function,  $E(Y | x)$ , and sketch a graph of this.

- d. Find  $E(XY)$  by the “double expectation” theorem, i.e.,

$$E(XY) = E[E(XY | X)] = E[X E(Y | X)] \text{ etc.}$$

Calculate the correlation coefficient between  $X$  and  $Y$ .

## Problem 2

- a.** Let  $Z$  be chi-square distributed with  $r$  degrees of freedom (written in short  $Z \sim \chi^2(r)$ ). According to the textbook, this is the same as the gamma distribution with parameters  $\alpha = r/2$  and  $\lambda = 1/2$  (in short  $Z \sim \Gamma\left(\frac{r}{2}, \frac{1}{2}\right)$ ). Use this and known properties of the gamma distribution to show that

(i)  $E(Z) = r, \quad \text{Var}(Z) = 2r$

- (ii) For  $r = 2$ , find the median of  $Z$ . Is the median smaller or larger than  $E(Z)$ ?

- b.** According to the textbook, if  $X$  is standard normally distributed ( $X \sim N(0, 1)$ ), then  $X^2 \sim \chi^2(1)$ .

- (i) Use this and known properties of the gamma distribution to show that, if

$X_1, X_2, \dots, X_n$  are iid with  $X_i \sim N(0, 1)$ , then  $Z = \sum_{i=1}^n X_i^2 \sim \chi^2(n)$ . (Note that iid means “independent and identically distributed”.)

- (ii) Use for example the central limit theorem to justify that  $Z = \sum_{i=1}^n X_i^2$  is approximately  $N(n, 2n)$  distributed for large  $n$ .

- c.** Let  $q_n$  denote the 95% quantile (i.e., such that  $P(Z \leq q_n) = 0.95$ ) in the  $\chi^2(n)$  distribution. Exact values for  $q_n$  can be found in the  $\chi^2$ -table in Rice. Let  $q_n^*$  denote the approximate 95% quantile determined by the approximate distribution for  $Z$  derived in section **b.** (i.e.,  $N(n, 2n)$ ). Show that

$$q_n^* = n + \sqrt{2n} \cdot 1.64$$

Calculate the approximation error  $q_n - q_n^*$  for  $n = 30, 60, 120$ , and comment on the result.

- d. The approximate quantile in section c. can be improved somewhat. Let  $X_1, X_2, \dots, X_n \sim N(0, 1)$  as in section b., and put

$$\bar{Z} = \frac{1}{n} \sum_{i=1}^n X_i^2$$

Then, by the central limit theorem combined with some other theory, it can be proven that  $\sqrt{2n}(\sqrt{\bar{Z}} - 1) \xrightarrow{D} N(0, 1)$  as  $n \rightarrow \infty$  (you do not need to justify this here). Show that, based on the approximate normal distribution for  $\sqrt{\bar{Z}}$ , we can derive an alternative approximation,  $q_n^{**}$ , to  $q_n$  (i.e., the 95% quantile of  $Z = n\bar{Z}$ ), given by

$$q_n^{**} = \frac{1}{2} \left( 1.64 + \sqrt{2n} \right)^2. \quad \text{Compare } q_n^{**} \text{ with } q_n^* \text{ for } n = 30, 60, 120.$$

[Note that this approximation is similar but slightly different from the suggestion given in the  $\chi^2$ -table in Rice, which represents a third approximation.]

### Problem 3

Suppose  $X_1, X_2, \dots, X_n$  are iid with  $X_i \sim \Gamma(9, \lambda)$ , where  $\lambda > 0$  is unknown. We assume here that the shape parameter is known,  $\alpha = 9$ .

- (i) Derive the maximum likelihood estimator for  $\lambda$  based on  $X_1, X_2, \dots, X_n$ .
- (ii) Develop an approximate 95% confidence interval for  $\lambda$ , and calculate the interval for the data

$$n = 100 \quad \text{and} \quad \sum_{i=1}^n X_i = 630$$