Econ 4130 UTSATT - Eksamen 2008 H

Sketch answers in << ….. >>.

Problem 1

a. Suppose $Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed (*iid*) random variables (rv's) where Y_i is Poisson distributed with parameter λ. Show that both the moment method estimator and the maximum likelihood estimator (*mle*) for λ are equal to the mean

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

<< Mme is trivial since $E(Y_i) = \lambda$.
The log likelihood is

$$l(\lambda) = -\sum_i \ln(y_i!) + \lambda \sum_i y_i - n\lambda \quad \text{and} \quad l'(\lambda) = \frac{1}{\lambda}\sum_i y_i - n = 0$$

gives $\hat{\lambda} = \bar{Y}$. >>

b. Show that the central limit theorem (applied to $\bar{Y}$) and the large sample theory for maximum likelihood estimators both lead to the same approximate normal distribution (for large *n*) for the mle $\hat{\lambda} = \bar{Y}$.

<< **Answer.** The CLT gives $\bar{Y} \overset{appr}{\sim} N\big(E(Y_i), \mathrm{var}(Y_i)/n\big) = N(\lambda, \lambda/n)$

The large sample mle theory gives
$\bar{Y} \overset{appr}{\sim} N\big(\lambda, 1/(nI(\lambda))\big)$ where $I(\lambda)$ is the Fisher information for one observation.
$\ln(f(y \mid \lambda)) = y\ln(\lambda) - \lambda - \ln y! \quad \Rightarrow$
$\ln' f = y/\lambda - 1 \quad \text{and} \quad \ln'' f = -y/\lambda^2$
Hence

$$I(\lambda) = -E\frac{\partial^2 \ln f}{(\partial\lambda)^2} = E(Y/\lambda^2) = 1/\lambda$$, giving the same approximate distribution >>

c. Let Y_i be the number of traffic accidents at a certain intersection during week no. *i*, for $i = 1,2,\ldots,50$. As in question **a** and **b** we assume that $Y_1, Y_2, \ldots, Y_{50}$ are iid and Pois(λ) distributed. The observations are summarized in table 1.

Table 1

No. of accidents	No. of weeks
0	32
1	12
2	6
3 or more	0
Sum	50

(i) Set up and calculate an approximate 95% confidence interval for λ based on the given data.
(ii) Justify the confidence interval using asymptotic properties of the mle and Slutsky's lemma.

<< **Answer.** $\hat{\lambda} = (1 \cdot 12 + 2 \cdot 6)/50 = 24/50 = 0.48$

95% CI:

$$\hat{\lambda} \pm 1.96\sqrt{\frac{\hat{\lambda}}{n}} = 0.48 \pm (1.96)(0.098) = [0.288,\ 0.572]$$ >>

d. We want to test the assumption that Y_i is Poisson distributed. For this purpose we need to modify our model. We still assume that $Y_1, Y_2, \ldots, Y_{50}$ are iid, but Y_i may have a different distribution than Poisson. Construct a chi-square test for the null-hypothesis, H_0 : "Y_i is Poisson distributed", based on the three categories, "$Y_i = 0$", "$Y_i = 1$", and "$Y_i \geq 2$", and the data in table 1. Test H_0, using 5% level of significance, and formulate a conclusion.

<< **Answer.**
Using the mle $\hat{\lambda} = 0.48$, we get

Category	Observed, O_i	$p_i(\lambda)$	Estimated, $E_i = np_i(\hat{\lambda})$	$(O_i - E_i)^2 / E_i$
$Y_i = 0$	32	$e^{-\lambda}$	30.94	0.036
$Y_i = 1$	12	$\lambda e^{-\lambda}$	14.85	0.547
$Y_i \geq 2$	6	$1 - e^{-\lambda} - \lambda e^{-\lambda}$	4.21	0.761
Sum	50		50	1.345

Hence $Q = \sum_{i=1}^{3} \frac{(O_i - E_i)^2}{E_i} = 1.345$, degrees of freedom = 3-1-1 = 1.

Critical value: $\chi^2_{0.95;1} = 3.84$. Conclusion: Don't reject H_0. >>

e. Suppose that the data were given in a more limited way as in table 2

Table 2

No. of accidents	No. of weeks
0	32
1 or more	18
Sum	50

where we do not know exactly how many weeks had 1 accident, how many had 2 accidents, and so on. I.e., we only know that 18 weeks had at least 1 accident. Accepting the model in question **a**, find the mle of λ under these circumstances and calculate an approximate 95% confidence interval for λ.
[**Hint:** Notice that the number of weeks with 0 accidents is binomially distributed.]

<< **Answer.** If X is the number of weeks with 0 accidents, then $X \sim \text{Bin}(50, p)$, where $p = e^{-\lambda}$. The simplest solution is to recognize that the Poisson model, in this case, simply represents a reparametrization of the binomial model, so we can use the ml-theory for the binomial model and transform the mle for p to find the mle for λ - as well as the CI. I.e. mle $\hat{p} = X/50 = 32/50 = 0.64$, and $\hat{\lambda} = -\ln(X/50) = -\ln(0.64) = 0.446$

$\text{se}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{50}} = 0.068$, which gives 95% CI for p: $\hat{p} \pm 1.96 \cdot \text{se}(\hat{p}) = [0.507, 0.773]$.
Transformed this gives 95% CI for λ: $[-\ln(U), -\ln(L)] = [0.257, 0.679]$

Alternative solution based on ml-theory:

Fisher info per obs: $f(x \mid p(\lambda))$ Bernoulli gives

$$\ln f = x \ln p(\lambda) + (1-x)\ln(1-p(\lambda)) = -x\lambda + (1-x)\ln(1-e^{-\lambda})$$

$$\partial \ln f = -x + (1-x)\frac{e^{-\lambda}}{1-e^{-\lambda}} = -x + (1-x)(\frac{1}{1-e^{-\lambda}} - 1)$$

$$\partial^2 \ln f = -(1-x)\frac{e^{-\lambda}}{(1-e^{-\lambda})^2}$$

and

$$I(\lambda) = -E\partial^2 \ln f = \frac{e^{-\lambda}}{1-e^{-\lambda}} = \frac{p}{1-p}$$

Appr. 95% CI directly from the asymp. distribution

$$\hat{\lambda} \pm 1.96 \cdot \sqrt{\frac{1-\hat{p}}{n\hat{p}}} = 0.446 \pm 1.96 \cdot 0.106 = [0.238, 0.654]$$

>>

Problem 2

Consider a fisherman selling his catch of fish on the marketplace in the local village a fixed day every week. Let *X* denote the size of his catch a given week measured as a proportion of the storage capacity of his fishing boat. Let *Y* denote the amount of fish he succeeds to sell at the marketplace the same week, also measured as a proportion of the storage capacity of his boat. As proportions, *X* and *Y* must both vary between 0 and 1. In addition we must have $Y \leq X$ since the fisherman cannot sell more than he has caught.

Assume that *X* and *Y* are continuous rv's with joint density (pdf) given by

$$f(x,y) = \begin{cases} 3x & \text{for } 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

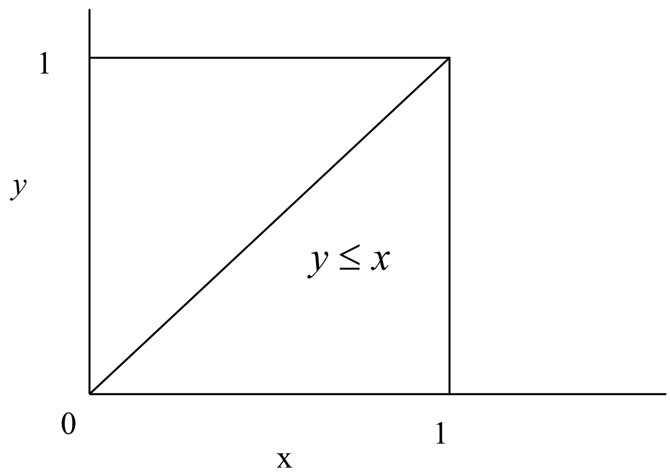
a. (**i**) Make a sketch of the region in the (x, y) - plane where the joint pdf $f(x,y) > 0$.

(**ii**) Show that the marginal pdf of *X* is given by

$$f_X(x) = \begin{cases} 3x^2 & \text{for } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(iii) Find $E(X)$ and $\text{var}(X)$.

<< **Answer.** (i)



(ii) $f_X(x) = \int_0^x 3x dy = 3x \int_0^x dy = 3x^2$ for $0 \le x \le 1$.

(iii) $E(X) = \int_0^1 3x^3 dx = \Big|_0^1 (3/4)x^4 = 3/4,$ $\quad E(X^2) = \int_0^1 3x^4 dx = \Big|_0^1 (3/5)x^5 = 3/5$

$\text{var}(X) = 3/5 - 9/16 = (48-45)/80 = 3/80$ >>

b. **(i)** Describe the conditional pdf, $f(y \mid x)$, for Y given $X = x$.
(ii) Find the conditional expectation, $E(Y \mid x)$, and variance, $\text{var}(Y \mid x)$, for Y given $X = x$.
(iii) Sketch a graph of the "regression function", $E(Y \mid x)$.

<< **Answer.** $f(y \mid x) = f(x,y)/f_X(x) = 3x/(3x^2) = 1/x$ for $0 < x < 1$. Hence $Y \mid x$ is uniform over $[0, x]$, which gives

$E(Y \mid x) = x/2$ and $\text{var}(Y \mid x) = x^2/12$.

Graph = …. >>

c. (**i**) Find $E(Y)$ and $\text{var}(Y)$. [**Hint:** Using the law of total expectation may save you some work.]
(**ii**) Find the correlation, $\text{corre}(X,Y)$, between X and Y. [**Hint**: Find first $E(XY) = E\left[X \cdot E(Y \mid X)\right] = \cdots$ etc.]

<< **Answer.** $E(Y) = E\left[E(Y \mid X)\right] = E(X/2) = 3/8$

$$\text{var}(Y) = E\left[\text{var}(Y \mid X)\right] + \text{var}\left[E(Y \mid X)\right] = E\left[X^2/12\right] + \text{var}\left[X/2\right] =$$

$$= \frac{3}{5 \cdot 12} + \frac{3}{4 \cdot 80} = \frac{3}{20}\left(\frac{1}{3} + \frac{1}{16}\right) = \frac{3(16+3)}{20 \cdot 48} = \frac{19}{320} = 0.0594$$

$$E(XY) = E\left[XE(Y \mid X)\right] = E\left[X^2/2\right] = (1/2)(3/5) = 3/10$$

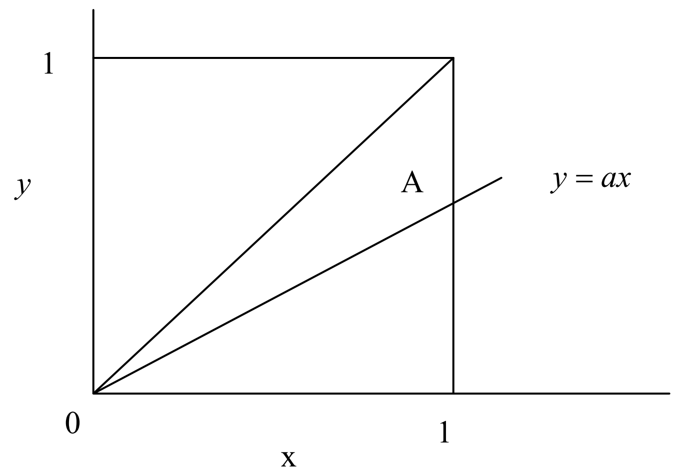
$$\text{cov}(X,Y) = \frac{3}{10} - \frac{3}{4} \cdot \frac{3}{8} = \frac{96-90}{320} = \frac{6}{320} = 0.01875$$

$$\text{corre}(X,Y) = \frac{6/320}{\sqrt{(12/320)(19/320)}} = \frac{6}{\sqrt{12 \cdot 19}} = \frac{6}{\sqrt{228}} = 0.397$$ >>

d. (**i**) Let a be a given positive number smaller than 1. Find $P(Y > aX)$ expressed by a.
(ii) What is the probability that the fisherman will succeed to sell at least 80% of his catch a given week?

<< **Answer.**

(i) see next page



We need to integrate *f* over the region marked A on the figure:

$$P(Y > aX) = \iint_A 3x\,dy\,dx = \int_0^1 \left(\int_{ax}^{x} 3x\,dy\right)dx = \int_0^1 3x(x - ax)\,dx = (1-a)\int_0^1 3x^2\,dx = 1 - a$$

(ii): Hence for a=0.80, $P(Y > 0.8X) = 0.20$. >>