TMA 4295 Statistical Inference 2014
Homework 8: Trial exam

Problem 1 (June 1993, Problem 1)

Let X have distribution Gamma(α, β), i.e., X has probability density function

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} \text{ for } x > 0,\ \alpha, \beta > 0$$

a) Show that the moment generating function (mgf) of X is given by

$$M_X(t) = (1 - \beta t)^{-\alpha} \text{ for } t < 1/\beta$$

Use $M_X(t)$ to find E(X) and Var(X).

Also, use the expression for $M_X(t)$ to find the mgf of $Z = \frac{2X}{\beta}$.

How can you conclude that $Z \sim \chi^2_{2\alpha}$? (Note that the degrees of freedom of a χ^2-distribution need not be an integer).

b) Show by using Chebyshev's inequality (Theorem 3.6.1) that

$$P(X \geq x) \leq \frac{M_X(t)}{e^{tx}}$$

for all $x > 0$ and $0 < t < 1/\beta$.

How would you choose t so that the upper bound is best possible?

Show that you then obtain the inequality

$$P(X \geq x) \leq \left(\frac{xe}{\alpha\beta}\right)^\alpha e^{-\frac{x}{\beta}} \qquad \text{for alle } x > \alpha\beta$$

c) It is known that X is approximately normally distributed when α is large and $\beta > 0$ is fixed.

Formulate and prove a precise result which expresses this by using convergence in distribution.

(*Hint:* You may study the mgf of $U = \frac{X - E(X)}{\sqrt{Var(X)}}$ when $\alpha \to \infty$.)

Problem 2 (May 1995, Problem 1)

In this problem you will need the definition of the multinomial distribution (see *Definition 4.6.2* in book):

A random vector $\underline{\mathbf{X}} = (X_1, \ldots, X_r)$ has a *multinomial distribution* with n trials, r cells and cell-probabilities $p_1, \ldots, p_r$ (where $0 \leq p_i \leq 1$ for $i = 1, \ldots, r$ and $\sum_{i=1}^r p_i = 1$) if the joint probability mass function for $\underline{\mathbf{X}}$ is given by

$$f(x_1, \ldots, x_r) = \frac{n!}{x_1! \cdots x_r!} p_1^{x_1} \cdots p_r^{x_r}$$

when $x_1, \ldots, x_r$ are nonnegative integers with $\sum_{i=1}^r x_i = n$.

The marginal distribution of X_i is binomial(n, p_i), $(i = 1, \ldots, r)$.

Animals of a specific kind can be categorized into one of four different types, depending on the occurrence or not of two factors. The probabilities that a randomly chosen animal is of type 1,2,3 or 4, respectively, are

$$p_1(\theta) = \frac{\theta+2}{4}, \quad p_2(\theta) = p_3(\theta) = \frac{1-\theta}{4}, \quad p_4(\theta) = \frac{\theta}{4}$$

where the parameter θ measures the degree of genetic coupling between the two factors, $(0 < \theta < 1)$.

To obtain information about the unknown parameter θ one observes a random sample of n animals, and lets X_j be the number of these that are categorized as type j $(j = 1, 2, 3, 4)$.

It is assumed that $\underline{\mathbf{X}} = (X_1, X_2, X_3, X_4)$ has a multinomial distribution with n trials, $r = 4$ cells and cell-probabilities $(p_1(\theta), p_2(\theta), p_3(\theta), p_4(\theta))$.

a) Show that the probability distribution of $\underline{\mathbf{X}}$ can be written as a two-parameter exponential family of the form

$$f(\underline{\mathbf{x}}|\theta) = h(\underline{\mathbf{x}})\theta^n \exp\{[\ln(\theta+2) - \ln(\theta)]x_1 + [\ln(1-\theta) - \ln(\theta)](x_2 + x_3)\}$$

What is the expression for $h(\underline{\mathbf{x}})$?

Explain why the statistic

$$T(\underline{\mathbf{X}}) = (X_1, X_2 + X_3)$$

is sufficient for θ.

Then prove that it is also *minimal sufficient.*

b) Derive an equation which the maximum likelihood estimator $\hat{\theta}$ for θ must satisfy.

Solve the equation and find $\hat{\theta}$ when $n = 200$ and $\underline{\mathbf{X}} = (125, 19, 21, 35)$.

c) Find an expression for the Cramér-Rao lower bound.

(*Hint:* It may be wise to use the formula in Lemma 7.3.11, which we have shown in class to be valid also for general joint distributions $f(\underline{\mathbf{x}}|\theta)$.)

d) Show that

$$\tilde{\theta}_1 = \frac{4}{n}X_1 - 2$$

as well as

$$\tilde{\theta}_2 = \frac{4}{n}(n - X_1 - X_2 - X_3) = \frac{4}{n}X_4$$

are unbiased estimators for θ.

Calculate the two estimators' variances and compare them. Compare also to the Cramér-Rao lower bound when $\theta = 0.5$.

Can Rao-Blackwell's theorem be used to improve on $\tilde{\theta}_1$ and/or $\tilde{\theta}_2$?

Prove, using $\tilde{\theta}_1$ and $\tilde{\theta}_2$, that $T(\underline{\mathbf{X}})$ is *not* complete.

Problem 3 (May 1995, Problem 2)

Let N be the number of accidents occurring at a specific company during a period of one year. The possible values of N are $0, 1, 2, \ldots$, and it is assumed that $\mathrm{E}(N)$ is finite. Let further $X_1, X_2, \ldots$ be the number of injured by the single accidents. It is assumed that the X_i are independent and identically distributed and in addition stochastically independent of N. Their common expectation is denoted by $\mathrm{E}(X)$ and is assumed to be finite.

The total number of injured during the year is now

$$S = \sum_{i=1}^{N} X_i$$

(where the empty sum $\sum_{i=1}^{0}$ is set to 0.)

(Note that the number of terms in the sum is here a random number N, not a fixed number n which you have seen most of the time earlier. It is the purpose of this exercise to study how such sums can be handled by proper conditioning on the values of N. The assumption that the X_i are independent of N is crucial in the following.)

a) Explain why

$$E(S|N = n) = nE(X) \qquad \text{for } n = 0, 1, \ldots$$

Then, show by the rule of "double expectation" that

$$\mathrm{E}(S) = \mathrm{E}(N)\mathrm{E}(X)$$

Let $M_N(t)$ and $M_X(t)$ denote the moment generating functions (mgf) of, respectively, N and the X_i. For simplicity we assume that these exist for all real t.

Show that the mgf of S is given by

$$M_S(t) = M_N(\ln M_X(t)) \qquad \text{for all } t$$

(*Hint:* First show that

$$\mathrm{E}(e^{tS}|N = n) = (M_X(t))^n \qquad \text{for } n = 0, 1, 2, \ldots)$$

b) Assume now that N is Poisson-distributed with parameter μ, and that the X_i are Poisson-distributed with parameter λ.

Show that the mgf of S in this case is

$$M_S(t) = e^{\mu(e^{\lambda(e^t-1)}-1)}$$

Use this to show that if $\mu \to \infty$ and $\lambda \to 0$ in such a way that $\lambda\mu \to a$ for a constant $a > 0$, then

$$S \xrightarrow{d} Y$$

where $Y \sim \text{Poisson}(a)$. (Here $\xrightarrow{d}$ means convergence in distribution).

How can this result be interpreted for practical use?