Università Commerciale Luigi Bocconi

MSc. Finance & CLEFIN A.A. 2013/2014

Prep Course in Statistics – Prof. Massimo Guidolin

A Few Review Questions and Problems concerning, Hypothesis Testing and Confidence Intervals

SUGGESTION: try to approach the questions first, without looking at the answers. Occasionally the questions/problems explicitly send you to check out concepts and definitions that have not explicitly discussed during the prep course. In this case—besides Casella and Berger's book—Wikipedia may be very useful, as always.

1. (Likelihood Ratio Test for a Normal Population Mean when the Variance is Unknown) Suppose X_1, ..., X_n is a random sample from a $N(\mu, \sigma^2)$ but an experimenter is interested only in inferences about μ, such as performing the one-sided mean test H_o: $\mu \leq \mu_0$ vs. H_1: $\mu > \mu_0$. In this case that the parameter σ^2 is a nuisance parameter. Derive the LRT and show that the test based on LRT($\mathbf{x}$) has the same structure as a test that defines the rejection region on the basis of a Student's t statistic.

2. (Computing and Optimizing a Normal Power Function) Let X_1, ..., X_n be a random sample from a $N(\mu, \sigma^2)$ population, σ^2 known. An LRT of H_o: $\mu \leq \mu_0$ versus H_1: $\mu.\mu_0$ is a test that rejects H_o if

$$\text{LRT} \equiv \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} > c.$$

The constant $c \in [0, 1]$ can be any positive number.

(a) Compute the power function of this LRT.

(b) Suppose the experimenter wishes to have a maximum Type I Error probability of 0.1. Suppose, in addition, the experimenter wishes to have a maximum Type II Error probability of 0.2 if $\mu \geq \mu_0 + \sigma$. Show how to choose c and n to achieve these goals.

3. (UMP Binomial Test Derived from Neyman-Pearson's Lemma) Notice that the conditions of Neyman-Pearson's lemma for simple hypotheses may be re-written as:

$$\begin{cases} \mathbf{x} \in R & \text{if } \frac{f(\mathbf{x};\theta_1)}{f(\mathbf{x};\theta_0)} > k \\ \mathbf{x} \in R^c & \text{if } \frac{f(\mathbf{x};\theta_1)}{f(\mathbf{x};\theta_0)} \leq k \end{cases} \quad \text{for some } k \geq 0 \text{ and } \alpha = \Pr_{\theta \in \Theta_o}(\mathbf{x} \in R),$$

assuming $f(\mathbf{x}; \theta_0) \neq 0$. Let $X \sim Bi(2, \theta)$. Use Neyman-Pearson lemma to show that in the test of H_o: $\theta = 1/2$ vs. H_1: $\theta = 3/4$, the test that rejects the null hypothesis if $X = 2$ is the UMP test at size $\alpha = 0.25$.

4. (Deriving Normal One-Sided Confidence Bounds by Inversion) Let X_1, ..., X_n be a random sample from a $N(\mu, \sigma^2)$ population. Construct a $1-\alpha$ upper confidence bound for μ, that is, we want a confidence interval of the form $C(\mathbf{x}) = (-\infty, U(\mathbf{x})]$.

5. (Confidence Interval for the Sample Variance) Let X_1, ..., X_n be a random sample from a $N(\mu, \sigma^2)$ population. Construct $1-\alpha$ confidence intervals for σ^2 and σ, respectively.

6. (MLEs and LRT-Based Tests in the Case of a Pareto Distribution) Let X_1, ..., X_n be a random sample from a Pareto population with PDF

$$f(x;\theta,v) = \frac{\theta v^\theta}{x^{\theta+1}} I_{[v,+\infty)}(x) \qquad \theta > 0,\ v > \theta \quad I_{[v,+\infty)}(x) = \begin{cases} 1 & \text{if } x \geq v \\ 0 & \text{otherwise} \end{cases}.$$

(a) Find the MLEs of θ and v.

(b) Show that the LRT of H_o: $\theta = 1$ and v unknown, versus H_1: $\theta \neq 1$ and v unknown, has critical region of the form $\{\mathbf{x} : \Lambda(\mathbf{x}) \leq c_1,\ \Lambda(\mathbf{x}) \geq c_2\}$, where $0 < c_1 < c_2$ and

$$\Lambda(\mathbf{x}) = \ln\left[\frac{\prod_{i=1}^n x_i}{(\min_i x_i)^n}\right] > 0$$

(*Hint*: note that by construction $\prod_{i=1}^n x_i/(\min_i x_i)^n > 1$).

7. (A Simple Application of the CLT) For a random sample X_1, ..., X_n of Bernoulli(p) variables, it is desired to test H_o: $p = 0.49$ versus H_1: $p = 0.51$. Use the Central Limit Theorem to determine, approximately, the sample size needed so that the two probabilities of error are both about .01. Use a test function that rejects H_o if $\sum_{i=1}^n x_i$ is large.

8. (Abusing the Size α in Practical Applications) One very striking abuse of α size levels is to choose them after seeing the data and to choose them in such a way as to force rejection (or acceptance) of a null hypothesis. To see what the true Type I and Type II error probabilities of such a procedure are, calculate size and power of the following two trivial tests:

(a) Always reject H_o, no matter what data are obtained (equivalent to the practice of choosing the α size level to force rejection of H_o).

(b) Always accept H_o, no matter what data are obtained (equivalent to the practice of choosing the α size level to force acceptance of H_o).