Exercises on nonparametric estimation of distribution functions and quantiles

Homework applied statistics

January 30, 2015

1. Data on the magnitudes of earthquakes near Fiji are available from R. Just type `quakes`. For help on this dataset type `?quakes`. Estimate the distribution function. Find an approximate 95% confidence interval for $F(4.9) - F(4.3)$ using a Wilson-type confidence interval.
   ([W], exercise 11, chapter 2.)

2. Hoeffding's inequality for Bernoulli random variables (equation 1.24 in [W]) can be used to construct a confidence interval for the success probability of a Binomial random variable with n trials. See example 1.17 of [W]. Assess the coverage probability of this confidence interval by means of a simulation study. How does this interval perform, compared to the Wilson interval (for example)?

3. Let $X_1, X_2, \ldots, X_n \sim F$ and let $\hat{F}_n(x)$ be the empirical distribution function. For a fixed X, find the limiting distribution of $\sqrt{F_n(x)}$. *Hint: use the Delta-method as explained in chapter 1 of [W].*
   ([W], exercise 4, chapter 2.)

4. A manufacturer wants to market a new brand of heat-resistant tiles which may be used on the space shuttle. A random sample of size m ot these tiles is put on a test and the heat resistance capacities of the tiles are measured. Let $X_{(1)}$ denote the smallest of these measurements. The manufacturer is interested in finding the probability that in a future test (performed by, say, an independent agency) of a random sample of n of these tiles at least k ($k = 1, 2, \ldots, n$) will have a heat resistance capacity exceeding $X_{(1)}$ units. Assume that the heat resistance capacities of these tiles follows a continuous distribution with cdf F.

   a. Show that the probability of interest is given by $\sum_{r=k}^{n} a_r$, where
   $$a_r = \frac{mn!(r+m-1)!}{r!(n+m)!} \; .$$

   b. Show that
   $$\frac{a_r}{a_{r-1}} = \frac{r+m-1}{r}$$

a relationship that is useful in calculating a_r.

c. Show that the number of tiles, n, to be put on a future test such that all of the n measurements exceed $X_{(1)}$ with probability p is given by

$$n = \frac{m(1-p)}{p} \ .$$

([Gibbons and Chakraborti (1992)], exercise 2.31)

References

[Gibbons and Chakraborti (1992)] Dickinson Gibbons, J. and Chakraborti, S (1992) *Nonparametric statistical inference*, 3rd edition, Marcel Dekker.

[W] Wasserman, L. (2005) *All of nonparametric statistics*, Springer.