STAT 331 Homework 8

Due date: In class on Thursday, November 18, 2004

Instructor: Dr. Rudolf Riedi

33. (15 points) You are given a set of data by a customer: 19.6, 1.88, 12.7 11.9, 6.85, 12.8 4.58, 19.0, 2.77, 13.6, 8.31, 11.7, 11.4, 13.8, 8.5, 13.2, 18.6, 8.46, 5.31, 6.99.

    (a) Calculate the sample mean.

    (b) Calculate the sample variance.

    (c) Calculate the sample standard deviation.

    (d) Make a range estimate of the true mean at the 90% confidence level.

    (e) The customer then asks for 5% accuracy and 95% confidence. Would 200 samples suffice to achieve this?

34. (10 points)

    For an exponential distribution with parameter λ, the mean is $1/\lambda$. Given are n samples t_i. An estimator for λ is proposed as

    $$\widehat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{t_i}.$$

    Is this estimator consistent? (No. Explain!)

35. (10 points)

    An experiment consists of a series of n Bernoulli trials with success probability p. The experiment is performed m times, and k_1, k_2, ..., k_m are the number of successes which were observed. An estimator is proposed for p as follows.

    For each experiment of n trials we let $\widehat{p}_i = k_i/n$ be an estimator for the success probability. The final estimator for p is then $\widehat{p} = 1/m \sum_{i=1}^{m} \widehat{p}_i$.

    (a) Is $\widehat{p}$ biased?

    (b) Is $\widehat{p}$ consistent?

    (c) Someone suggests to consider all mn Bernoulli trials as one experiment and to use the overall total number of successes over the number of trials as an estimator, i.e., $\widehat{\pi} = 1/(mn) \sum_{i=1}^{m} k_i$. Does the performance of the estimation improve or not? (Compare the bias and variance.)

36. (5 points)

    Compute the Maximum Likelihood Estimator (MLE) of the Rayleigh distribution. Here is the procedure. Let

    $$f(r) = \frac{r}{\sigma^2} \exp\left(-\frac{1}{2}\left(\frac{r}{\sigma}\right)^2\right) \qquad r > 0,\ \text{zow}$$

    denote the Rayleigh density. The parameter to estimate is σ—or σ^2—whichever is easier to obtain. For simplicity, we write $b = 2\sigma^2$ and actually estimate b.

    Assume n independent samples of this distribution are given. Their joint density is then the n-fold product of f which we denote by $f_n(r_1, ..., r_n) = f(r_1) \cdots f(r_n)$. Now, the value of b which maximizes f_n is the MLE which we denote by $\widehat{b}$. (Hint: Quite often in MLE computation it is convenient to maximize $\log(f_n)$. Since $\log(\cdot)$ is a monotonous function, this is equivalent to maximizing f_n itself.)