Problem Set 10

Due: Wednesday April 30, 2008

Note: We will refer to the text as [YG].

Readings: Course Notes for Lecture 19-20 and corresponding parts of [YG], Chapters 6 & 7.

Problem 1

Let X be an r.v. with mean m and variance σ^2. Given two random samples of size n_1 and n_2 with sample means $\bar{X}_1$ and $\bar{X}_2$, respectively, show that

$$\bar{X} = a\bar{X}_1 + (1-a)\bar{X}_2, \qquad 0 < a < 1$$

is an unbiased estimator of m. Assuming $\bar{X}_1$ and $\bar{X}_2$ to be independent, find the value of a that minimizes the variance of $\bar{X}$.

Problem 2

We are interested in the arrival rate of customers that use the BankBoston ATM near Warren towers during lunchtime, i.e., during the period 12:00pm-1:00pm. Assume that the arrivals to the ATM follow a Poisson process with rate λ per hour. Then, the number of arrivals during lunchtime is a Poisson r.v., which we will denote by X. Suppose that we measure X over n days. That is, we observe $X_1, X_2, \dots, X_n$ which are independent Poisson r.v.'s with parameter λ. Let M_n denote the sample mean, i.e.,

$$M_n = \frac{\sum_{i=1}^{n} X_i}{n}.$$

(a) Find $\mathbf{E}[M_n]$ and $\mathbf{V}[M_n]$.

(b) Find a 95% confidence interval for λ.
Hint: Consider the statistic $(M_n - \lambda)/\sqrt{\lambda/n}$.

Problem 3

The life in hours of a 75-watt light bulb is known to be approximately normally distributed with $\sigma = 25$ hours. A random sample of 20 bulbs has a mean life of 1014 hours. Construct a 99% two-sided confidence interval on the mean life.

Problem 4

The return of the Dow Jones stockmarket index over the last 35 years is given by the following data

10.3818,	10.3783,	7.9247,	8.6546,	8.3493,	7.6266,	9.4516,
6.8234,	12.3664,	7.7272,	8.2279,	10.1335,	8.1186,	7.8087,
6.3353,	8.5888,	5.3276,	9.4286,	11.2471,	6.6164,	9.7160,
10.5080,	4.8125,	5.1181,	9.1423,	7.2002,	9.3800,	9.6312,
9.4238,	10.5805,	9.3372,	10.3817,	5.5951,	7.9604,	7.6866

Construct a 95% confidence interval on the mean annual return.

Problem 5

Consider two independent r.v.'s X_1 and X_2 with unknown means m_1, m_2 and known variances σ_1^2, σ_2^2, respectively. Let $x_{1,1}, x_{1,2}, \ldots, x_{1,n_1}$ be a random sample of n_1 observations from X_1 and $x_{2,1}, x_{2,2}, \ldots, x_{2,n_2}$ a random sample of n_2 observations from X_2. Assume that n_1 and n_2 are fairly large. Construct an $(1-a)100\%$ confidence interval on the difference of the two means $m_1 - m_2$.

Problem 6

The Dean of Engineering at BU keeps track of the average GPA overall graduating Engineering students. That is, assuming that there are n students graduating on a particular year, the Dean computes

$$M_n = \frac{\sum_{i=1}^{n} G_i}{n},$$

where G_i is the GPA of student $i = 1, \ldots, n$ graduating during that year. Suppose that $n = 80$ students graduate every year and that the standard deviation of a student's GPA is 0.8. The Dean has kept M_n for the last 20 years (in order):

$$3.0691, 2.9839, 3.1905, 3.2759, 3.1339, 3.1000, 3.1928, 3.1650, 3.0085, 3.4585,$$
$$3.2882, 3.4562, 3.1005, 2.9763, 3.3615, 3.2082, 3.0488, 3.1822, 2.7982, 3.4168$$

Plot a control chart for M_n. Is the "grading process" in control? If not which years are problematic?