Problem set 4 – Submit your homework to the box called “Probability and Statistics Uluğ Bayazıt” in Computer Engineering Secretary’s Office until May 12, 2016. Late submissions will not be graded..

The solutions will be posted on Ninova before the first week of final exams.

1. X_1 and X_2 are random variables that satisfy $E[X_1] = E[X_2] = \mu$ ve $Var[X_1] = 10$, $Var[X_2] = 15$.

a) Determine the bias and variance of the point estimate $\hat{\mu}_1 = \dfrac{X_1 + 2X_2}{6} + 9$

b) Specify the range for μ so that the point estimate $\hat{\mu}_1$ has a a lower mean squared error than the point estimate $\hat{\mu}_2 = \dfrac{X_1 + X_2}{2}$?

2. X is a binomial random variable defined over 12 Bernoulli trials with a success probability of p in each (i.e. $X \sim B(12, p)$). Consider $\hat{p} = \dfrac{X}{10}$

a) What is the bias of this point estimate?

b) What is the variance of this point estimate?

c) Write down the approximate sampling distribution of the point estimate. (Note that binomial distribution converges to Normal distribution, but binomial density does not converge to Normal density)

d) Determine the mean squared error of this point estimate. Is the mean squared error of this estimate always worse than the mean squared error of the unbiased point estimate $\hat{p} = \dfrac{X}{12}$

3. Let a point estimate for the sample variance be given as $\hat{\sigma}^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ where n is the number of samples.

a)What is the bias in this estimate as a function of n ?

b)How much can you change the mean squared error of this estimate if you use the point estimate for the sample variance that gives the lowest bias? (Hint: Take $Var(\hat{\sigma}^2) = \dfrac{\sigma^4(n-1)}{n^2}$)

4. In a consumer survey, 250 people out of a representative sample of 450 people say that they prefer product A to product B. Let *p* be the proportion of all consumers who prefer product A to product B. Construct a point estimate of *p*. What is the standard error of your point estimate? What is the mean squared error of your point estimate?

5. An atlete specialized in long jump events jumps an average of $\bar{x} = 7.91m$ in 15 trials. The standard error of the mean jump distance in these trials is 0.2m.

i) Determine a two sided confidence interval around the mean estimate that includes the mean jump distance of the atlete with a confidence level of 95%.

ii) Determine a one sided confidence interval around the mean estimate that includes the mean jump distance of the atlete with a confidence level of 90% and bounds the estimate only from below. Explain why one might want to use a one sided confidence interval like this one than a two sided one as in i)

iii) If the atlete performs 500 jumps, how likely is it that the average distance of these jumps is 8.05m or more?

6. The densities of 100 solutions in chemical solution bottles are measured. The sample mean turns out to be $\bar{x} = 0.7680$ and the sample standard deviation is $\hat{\sigma} = 0.042$. Determine the constant c so that $\mu \in (c, \infty)$ is a one-sided confidence interval with 95% confidence level. Is it likely that the mean density is less than 0.7675?

7. The density of the lifetimes of lightbulbs manufactured by a certain company is known to be distributed as Normal with a known standard deviation of 100 hours. A sample of 8 lightbulbs have lifetimes {1552hours, 1475 hours, 1505 hours, 1462 hours, 1531hours,1538 hours, 1511 hours, 1430 hours}. Determine the sample mean distribution. Determine c so that the one sided confidence interval (c, ∞) includes the true mean with a confidence level of 95%. Explain why one might want to use a one sided confidence interval of the form (c, ∞) than a one sided confidence interval of the form $(-\infty, c)$

8. Screws manufactured in a factory come from a Normal distribution. A sample of 9 screws have a mean length of $\bar{x} = 2.55cm$ and a sample standard deviation of $\hat{\sigma} = 0.1cm$. Test the hypothesis that $\mu = 2.6cm$. Assume that the test has a significance level of $\alpha = 0.05$.

9. A quality control inspector would like to determine whether the mean thickness value for glass panels is 3.0mm or not by applying hypothesis testing. Take $H_0 : \mu = 3.0mm$, $H_A : \mu \neq 3.0mm$. A sample of 41 glass panels are examined.

a) At a test significance level of $\alpha = 0.05$, for which values of the t-statistic should H_0 be accepted.

b) At a test significance level of $\alpha = 0.01$, for which values of the t-statistic should H_0 be rejected.

c) If the sample mean is $\bar{x} = 3.04mm$ and the sample standard deviation is $\hat{\sigma} = 0.06mm$ should H_0 be accepted for a test significance level of $\alpha = 0.1$? Repeat for $\alpha = 0.01$. Determine the p-value.

10. Let $H_0 : \mu \leq 3.0mm$ in the previous problem. Repeat the previous problem for this hypothesis.

11. In problem 9 a)suppose you accept H_0.for a particular sample of 41 glass panels. When you construct a confidence interval with 95% confidence level around this sample mean observation, does the confidence interval necessarily include 3.0mm? Explain.

12. A sample of 28 metal cylinder thickness measurements has a sample mean of 3.28 milimeters. If an experimenter wishes to use a “known” value of $\sigma = 1.5$ milimeters for the population standard deviation, construct an appropriate 95% two-sided confidence interval for the population mean μ. Test the hypothesis that $\mu = 4.00$ milimeters for significance level of $\alpha = 0.05$.