HW 2: Stat 430

[Chapter 2: Point estimation and Confidence Intervals]

(a) **Biased-Unbiased Parameter Estimation** Show that for n i.i.d values $X_1, \ldots, X_n$ with $E[X_i] = \mu$ and $Var[X_i] = \sigma^2$

- $s^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$ is a biased parameter estimate for σ^2.
- $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ is an unbiased parameter estimate for σ^2.

Hint:

The sum $\sum_{i=1}^{n} (X_i - \bar{X})^2$ can be written as $(\sum_{i=1}^{n} X_i^2) - n\bar{X}^2$.

If Y is a random variable, the expected value of Y^2 can be computed as: $E[Y^2] = Var[Y] + (E[Y])^2$

(b) **Method of Moments**

The Pareto distribution has two parameters $a > 0$ and $b > 0$, its density is given as:

$$f_{a,b}(x) = \begin{cases} \frac{ab^a}{x^{a+1}} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

(i) Use the Method of Momentss to get estimates for the parameters a and b, i.e. compute the first and second moments of the Pareto distribution. This will give you two equations for a and b. Solve for a and b then to get $\hat{a}$ and $\hat{b}$.

(ii) Assume that the following ten numbers come from a Pareto distribution with unknown parameters a and b. Use the Method of Moments to get estimates for a and b:

$$x_1 = 1.30, x_2 = 33.92, x_3 = 1.67, x_4 = 1.81, x_5 = 22.30, x_6 = 4.62, x_7 = 2.22, x_8 = 20.73, x_9 = 2.52, x_{10} = 17.96$$

(c) **Google - search response times**

The following numbers represent the response times (in s) of seventeen queries to Google's search machine (www.google.com):

$$0.21, 0.15, 0.09, 0.06, 0.14, 0.1, 0.1, 0.12, 0.06, 0.13, 0.24, 0.15, 0.38, 0.08, 0.14, 0.1, 0.12.$$

(i) Find mean, median, mode, minimum and maximum of the observed response times.

(ii) Assume, the above values are i.i.d exponentially distributed with some (unknown) rate λ. Derive an ML-Estimator for $1/\lambda$ and give an estimate $\widehat{1/\lambda}$ for the above values.

(d) **Highway Speed**

There is concern about the speed of automobiles traveling over a particular stretch of highway. For a random sample of thirty automobiles, radar indicated the following speeds, in miles per hour:

82 78 64 78 77 57 74 70 81 80 75 78 85 77 78 69 73 79 78 66 71 70 61 65 66 57 72 67 64 74

(i) Find the sample mean and variance.

(ii) Find a 95 % confidence interval for the mean speed of all automobiles travelling over this stretch of highway.

(e) **Laboratory Scale**

To assess whether a laboratory scale is accurate we can take a standard weight known to weigh exactly 100 grams and weigh it repeatedly. Suppose that the scale readings are normally distributed with standard deviation $\sigma = 0.25$ grams. If the scale is accurate then the population mean μ(the mean obtained in many repeated weighings) would be 100 grams but if the scale is inaccurate the population mean could be higher or lower.

(i) The weight is weighed 35 times and the sample mean is $\bar{X} = 99.90$. Construct a 99% confidence interval for μ. Do you believe the scale is accurate based on this interval?

(ii) Construct a 90% confidence interval for μ. Do you believe the scale is accurate based on this interval?

(iii) Explain why one confidence interval finds the scale inaccurate while the other finds the scale accurate.

(iv) The analysis is criticized because it turns out that the scale's measurements are not approximately normal. Explain why this is not really a problem.

(f) **Program Execution Times**

A program was tested on 30 data sets, execution times were measured. Sample mean and deviation for execution times are: $\bar{X} = 65$ ms and $s = 6$ ms.

Compute a 90% and a 95% confidence interval for the mean response time.