# HUL315 Assignment-2

Amaiya Singhal 2021CS50598

## Problem 1

**a)** Verify that if $X_1, ..., X_n$ are a random sample drawn from a $\mu, \sigma^2$ distribution, then $z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ is $N(0, 1)$.

*Proof.* We use the Central Limit theorem.

> The **Central Limit Theorem** states that, if $X_1, X_2, ..., X_n$ is a random sample from a distribution with mean $\mu$ and variance $\sigma^2$ , then the distribution of $\bar{X}$ tends to the normal distribution $N(\mu, \sigma^2/n)$ as $n \to \infty$

We can also verify that the mean and variance are correct.

$$\bar{X} = \frac{X_1 + X_2 + ... + X_n}{n}$$

$$E[\bar{X}] = E\left[\frac{X_1 + X_2 + ... + X_n}{n}\right]$$

$$E[\bar{X}] = \frac{E[X_1 + X_2 + ... + X_n]}{n}$$

Now, using linearity of expectation,

$$E[\bar{X}] = \frac{E[X_1] + E[X_2] + ... + E[X_n]}{n}$$

$$E[\bar{X}] = \frac{\mu + \mu + ... + \mu}{n}$$

$$E[\bar{X}] = \frac{n\mu}{n}$$

$$E[\bar{X}] = \mu$$

$$\bar{X} = \frac{X_1 + X_2 + ... + X_n}{n}$$

$$Var[\bar{X}] = Var\left[\frac{X_1 + X_2 + ... + X_n}{n}\right]$$

$$Var[\bar{X}] = \frac{Var[X_1 + X_2 + ... + X_n]}{n^2}$$

As the random samples are i.i.d.,

$$Var[\bar{X}] = \frac{Var[X_1] + Var[X_2] + ... + Var[X_n]}{n^2}$$

$$Var[\bar{X}] = \frac{\sigma^2 + \sigma^2 + ... + \sigma^2}{n}$$

$$Var[\bar{X}] = \frac{n\sigma^2}{n^2}$$

$$Var[\bar{X}] = \frac{\sigma^2}{n}$$

□

From the Central Limit Theorem, we can conclude that $\bar{X} \sim N(\mu, \sigma^2/n)$

The Cumulative distribution function for the normal distribution is:

$$P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}\, dy$$

For $\bar{X}$, the CDF is

$$P(\bar{X} \leq x) = \int_{-\infty}^{x} \frac{1}{(\sigma/\sqrt{n})\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2(\sigma/\sqrt{n})^2}}\, dy$$

We define a new random variable $Z$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

$$E[Z] = E\left[\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right]$$

$$E[Z] = \frac{E[\bar{X} - \mu]}{\sigma/\sqrt{n}}$$

$$E[Z] = \frac{E[\bar{X}] - \mu}{\sigma/\sqrt{n}}$$

$$E[Z] = \frac{\mu - \mu}{\sigma/\sqrt{n}}$$

$$E[Z] = 0$$

$$Var[Z] = Var\left[\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right]$$

$$Var[Z] = \frac{Var[\bar{X} - \mu]}{\sigma^2/n}$$

$$Var[Z] = \frac{Var[\bar{X}]}{\sigma^2/n}$$

$$Var[Z] = \frac{\sigma^2/n}{\sigma^2/n}$$

$$Var[Z] = 1$$

Using the CDF of $\bar{X}$ we can write

$$P(\bar{X} \leq (z(\sigma/\sqrt{n}) + \mu)) = \int_{-\infty}^{z(\sigma/\sqrt{n})+\mu} \frac{1}{(\sigma/\sqrt{n})\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2(\sigma/\sqrt{n})^2}}\, dy$$

substituting $y = x(\sigma/\sqrt{n}) + \mu$

$$P\left(\bar{X} \leq (z(\sigma/\sqrt{n}) + \mu)\right) = \int_{-\infty}^{z} \frac{1}{(\sigma/\sqrt{n})\sqrt{2\pi}} e^{-\frac{(x(\sigma/\sqrt{n})+\mu-\mu)^2}{2(\sigma/\sqrt{n})^2}}\, (\sigma/\sqrt{n})dx$$

$$P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z\right) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

$$P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

which is the CDF of the standard normal distribution. Hence, we can conclude $Z \sim N(0, 1)$

**b)** Use the fact that $(n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$ to show that $t = z/\sqrt{s^2/\sigma^2} = (\bar{X}-\mu)/(s/\sqrt{n})$ has a t-distribution with $(n-1)$ degrees of freedom.

*Proof.*

$$t = \frac{\bar{X}-\mu}{s/\sqrt{n}}$$

$$t = \frac{\bar{X}-\mu}{s/\sqrt{n}}.\frac{\sigma/\sqrt{n}}{\sigma/\sqrt{n}} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}.\frac{\sigma/\sqrt{n}}{s/\sqrt{n}}$$

Let $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ which follows the standard normal distribution (proved earlier)

$$t = Z.\frac{\sigma}{s}$$

$$t = \frac{Z}{s/\sigma}$$

$$t = \frac{Z}{\sqrt{s^2/\sigma^2}}$$

$$t = \frac{Z}{\sqrt{(n-1)s^2/(n-1)\sigma^2}}$$

$$t = \frac{Z}{\sqrt{\frac{(n-1)s^2/\sigma^2}{n-1}}}$$

Let $X = (n-1)s^2/\sigma^2$ which follows a chi-square distribution, $X \sim \chi^2_{n-1}$ (given). On substitution,

$$t = \frac{Z}{\sqrt{X/(n-1)}}$$

where the numerator $Z \sim N(0,1)$ and the denominator is the square root of $X \sim \chi^2_{n-1}$ divided by its degree of freedom $n-1$. Hence, $t \sim t_{n-1}$

$\square$

**c)** For $n = 16$, $\bar{x} = 20$ and $s^2 = 4$, construct a 95% confidence interval for $\mu$.

*Solution.* From the previous question we get that $(\bar{X}-\mu)/(s/\sqrt{n})$ follows the t-distribution. Now, using the t-distribution table we get

$$P\left(-2.131 < \frac{\bar{X}-\mu}{s/\sqrt{n}} < 2.131\right) = 0.95$$

$$P\left(\bar{X} - 2.131\frac{s}{\sqrt{n}} < \mu < \bar{X} + 2.131\frac{s}{\sqrt{n}}\right) = 0.95$$

$$P\left(20 - 2.131\frac{2}{\sqrt{16}} < \mu < 20 + 2.131\frac{2}{\sqrt{16}}\right) = 0.95$$

$$P\left(20 - 2.131\frac{2}{4} < \mu < 20 + 2.131\frac{2}{4}\right) = 0.95$$

$$P\left(20 - 1.0655 < \mu < 20 + 1.0655\right) = 0.95$$

$$P\left(18.9345 < \mu < 21.0655\right) = 0.95$$

Hence, the 95% confidence interval for $\mu$ is $[18.9345, 21.0655]$

## Problem 2

Let $\bar{Y}$ denote the sample average from a random sample with mean $\mu$ and variance $\sigma^2$. Consider two alternative estimators of $\mu$ : $W_1 = [(n-1)/n]\bar{Y}$ and $W_2 = \bar{Y}/2$.

**a)** Show that $W_1$ and $W_2$ are both biased estimators of $\mu$ and find the biases. What happens to the biases as $n \to \infty$?

*Proof.* To check that $W_1$ and $W_2$ are biased we need to calculate $E[W_1]$ and $E[W_2]$

> **Unbiasedness**: An estimator, $W$ of $\theta$, is an unbiased estimator if $E(W) = \theta$.
> **Bias** is defined as Bias $(W) = E(W) - \theta$

$$E[W_1] = E\left[\left(\frac{n-1}{n}\right)\bar{Y}\right]$$

$$E[W_1] = \left(\frac{n-1}{n}\right)E[\bar{Y}]$$

$$E[W_1] = \left(\frac{n-1}{n}\right).\mu \neq \mu$$

Since $E[W_1] \neq \mu,\ W_1$ is a biased estimator of $\mu$.

$$\text{Bias}[W_1] = E[W_1] - \mu$$

$$\text{Bias}[W_1] = \left(\frac{n-1}{n}\right).\mu - \mu$$

$$\text{Bias}[W_1] = \mu\left(\frac{n-1}{n} - 1\right)$$

$$\text{Bias}[W_1] = \mu\left(\frac{n-1-n}{n}\right) = -\frac{\mu}{n}$$

$$E[W_2] = E\left[\frac{\bar{Y}}{2}\right]$$

$$E[W_2] = \frac{E[\bar{Y}]}{2}$$

$$E[W_2] = \frac{\mu}{2} \neq \mu$$

Since $E[W_2] \neq \mu,\ W_2$ is a biased estimator of $\mu$.

$$\text{Bias}[W_2] = E[W_2] - \mu$$

$$\text{Bias}[W_2] = \frac{\mu}{2} - \mu$$

$$\text{Bias}[W_2] = -\frac{\mu}{2}$$

Considering the case when $n \to \infty$

$$\text{Bias}[W_1] = \lim_{n\to\infty} -\frac{\mu}{n} = 0$$

$$\text{Bias}[W_2] = \lim_{n\to\infty} -\frac{\mu}{2} = -\frac{\mu}{2}$$

□

**b)** Find the probability limits of $W_1$ and $W_2$. Which estimator is consistent?

*Solution.*

$$\text{plim}[W_1] = \lim_{n\to\infty}\left(\frac{n-1}{n}\right)\bar{Y} = \lim_{n\to\infty}\left(1 - \frac{1}{n}\right)\bar{Y} = \mu$$

$$\text{plim}[W_2] = \lim_{n\to\infty}\frac{\bar{Y}}{2} = \frac{\mu}{2} \neq \mu$$

$W_1$ is a consistent estimator of $\mu$ as $plim(W_1) = \mu$, however $W_2$ is inconsistent because $plim(W_2) = \mu/2 \neq \mu$

## Problem 3

Suppose that you have two independent unbiased estimators of the same parameter $\alpha$, say $\hat{a_1}$ and $\hat{a_2}$, with different standard deviations $\sigma_1$ and $\sigma_2$. What linear combination of $\hat{a_1}$ and $\hat{a_2}$ is the minimum variance unbiased estimator of $\alpha$?

*Solution.* Let us denote the minimum variance estimator as $\hat{a}_{min} = c_1\hat{a_1} + c_2\hat{a_2}$.

As the estimator is unbiased: $E[\hat{a}_{min}] = E[\hat{a_1}] = E[\hat{a_2}]$

$$E[\hat{a}_{min}] = E[c_1\hat{a_1} + c_2\hat{a_2}]$$

$$E[\hat{a}_{min}] = E[c_1\hat{a_1}] + E[c_2\hat{a_2}]$$

$$E[\hat{a}_{min}] = c_1E[\hat{a_1}] + c_2E[\hat{a_2}]$$

$$E[\hat{a}_{min}] = (c_1 + c_2)E[\hat{a}_{min}]$$

$$c_1 + c_2 = 1$$

We can now rewrite $\hat{a}_{min}$ as $c_1\hat{a_1} + (1 - c_1)\hat{a_2}$

$$Var[\hat{a}_{min}] = Var[c_1\hat{a_1} + (1 - c_1)\hat{a_2}]$$

$$Var[\hat{a}_{min}] = Var[c_1\hat{a_1}] + Var[(1 - c_1)\hat{a_2}]$$

$$Var[\hat{a}_{min}] = c_1^2Var[\hat{a_1}] + (1 - c_1)^2Var[\hat{a_2}]$$

$$Var[\hat{a}_{min}] = c_1^2\sigma_1^2 + (1 - c_1)^2\sigma_2^2$$

For finding out the value of $c_1$ for which $Var[\hat{a}_{min}]$ is the least, we impose the conditions

$$\frac{\partial Var[\hat{a}_{min}]}{\partial c_1} = 0, \quad \frac{\partial^2 Var[\hat{a}_{min}]}{\partial^2 c_1} > 0$$

$$\frac{\partial Var[\hat{a}_{min}]}{\partial c_1} = 2c_1\sigma_1^2 - 2(1 - c_1)\sigma_2^2 = 0$$

$$(\sigma_1^2 + \sigma_2^2)c_1 - \sigma_2^2 = 0$$

$$c_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

$$\frac{\partial^2 Var[\hat{a}_{min}]}{\partial^2 c_1} = 2(\sigma_1^2 + \sigma_2^2) > 0$$

As variance is a positive quantity, we get that the double derivate is positive which tells us that the value of $c_1$ we have obtained gives us the minimum value of the expression. So we get $\hat{a}_{min}$ as

$$\hat{a}_{min} = \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right)\hat{a_1} + \left(\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)\hat{a_2}$$