EEL 4930 Stats – Lecture 27

EEL 4930 Stats – Lecture 27

- **Review of in-class assignment from last lecture**

EEL 4930 Stats – Lecture 27

- **Review of in-class assignment from last lecture**

  Recall that if x_i are samples drawn from a random variable X, then

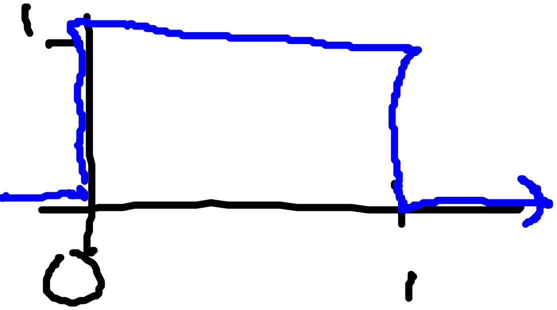
$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} x_i = E[X].$$

  Create a Uniform random variable object using scipy.stats. Draw 10,000 sample values from it, and use the sample values to estimate $(E[U])^2$ and $E[U^2]$.

Analytical Solution: $(E[U])^2$, $E[U^2]$

$U \sim$ uniform$[0,1]$

$$f_U(u) = \begin{cases} 1, & 0 \le u \le 1 \\ 0, & \text{o.w} \end{cases}$$



$$E[U] = \int_{-\infty}^{\infty} u \, f_U(u)\, du$$

$$= \int_0^1 u(1)\, du \rightarrow (\ )^2$$

$$E[U^2] = \int_{-\infty}^{\infty} u^2 f_U(u)\, du$$

By LOTUS

$$= \int_0^1 u^2 (1)\, du$$

- Find the value c that minimizes the expected mean-square error to a random variable X, $E[(X-c)^2]$

$$\frac{d}{dc} E\left[(X-c)^2\right] = 0$$

$$E\left[\frac{d}{dc}(X-c)^2\right] = 0 \quad \text{by linearity}$$

$$E\left[2(X-c)(-1)\right] = 0$$

$$\cancel{(-2)} E\left[X-c\right] = 0 \quad \text{by lin.}$$

$$E[X] - c = 0 \quad \text{" "}$$

$$c = E[X]$$

The Sample Mean Estimator

- Let $X_i, \quad i = 0, 1, \ldots, N-1$ be a finite sequence of independent random variables drawn from the same distribution, which has some finite mean μ.

THE SAMPLE MEAN ESTIMATOR

- Let $X_i, \quad i = 0, 1, \ldots, N-1$ be a finite sequence of independent random variables drawn from the same distribution, which has some finite mean μ. If we wish to estimate the mean of the underlying distribution, we can do it as in our last simulation:

$$\hat{\mu} = \frac{1}{N} \sum_{i=0}^{N-1} X_i$$

THE SAMPLE MEAN ESTIMATOR

- Let $X_i, \quad i = 0, 1, \ldots, N-1$ be a finite sequence of independent random variables drawn from the same distribution, which has some finite mean μ. If we wish to estimate the mean of the underlying distribution, we can do it as in our last simulation:

$$\hat{\mu} = \frac{1}{N} \sum_{i=0}^{N-1} X_i$$

- This is called the **sample mean estimator**

- Note that the sample mean estimator is a *linear combination* of random variables. Thus,

- Note that the sample mean estimator is a *linear combination* of random variables. Thus,

1. the sample mean estimator is a random variables

- Note that the sample mean estimator is a *linear combination* of random variables. Thus,

1. the sample mean estimator is a random variables
2. we can calculate it's expected value!

$$E[\hat{\mu}] = E\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right]$$

$$= \underbrace{\frac{1}{N}\sum_{i=1}^{N} \underbrace{E[X_i]}_{\mu}}_{N\mu}$$

$$= \frac{N\mu}{N} = \mu$$

Estimator is unbiased if E[ ] = thing being estimated

Moments

- Moments of a random variable are expected values of the random variable raised to some power

MOMENTS

- Moments of a random variable are expected values of the random variable raised to some power
- For a central moment, the mean is subtracted from the random variable before it is raised to a power

- Because different powers spread the values of the random variable in different ways, moments can provide additional information about a random variable than the mean

- Because different powers spread the values of the random variable in different ways, moments can provide additional information about a random variable than the mean
  - Variance is the second central moment and provides a measure of how much the probability density or mass of random variable is spread away from the mean

- Some common moments (expected values):

- Some common moments (expected values):
  - nth moment of X:

- Some common moments (expected values):
  - nth moment of X:

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx.$$

- Some common moments (expected values):
  - nth moment of X:

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx.$$

  - nth central moment of X:

- Some common moments (expected values):
  - nth moment of X:

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx.$$

  - nth central moment of X:

$$E[(X - \mu_X)^n]$$

- Some common moments (expected values):
  - nth moment of X:

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)dx.$$

  - nth central moment of X:

$$E[(X-\mu_X)^n] = \int_{-\infty}^{\infty} (x-\mu_X)^n f_X(x)dx,$$

where $\mu_X = E[X]$.

– Variance of X is 2nd central moment:

– Variance of X is 2nd central moment:

$$\text{Var}[X] \;=\; E[(X-\mu_X)^2]$$

– Variance of X is 2nd central moment:

$$\begin{aligned}
\text{Var}[X] &= E[(X-\mu_X)^2] \\
&= E[X^2 - 2X\mu_X + \mu_X^2]
\end{aligned}$$

– Variance of X is 2nd central moment:

$$\begin{aligned}
\text{Var}[X] &= E[(X-\mu_X)^2] \\
&= E[X^2 - 2X\mu_X + \mu_X^2] \\
&= E[X^2] - 2E[X]\mu_X + \mu_X^2
\end{aligned}$$

– Variance of X is 2nd central moment:

$$\begin{aligned}
\mathrm{Var}[X] &= E[(X-\mu_X)^2] \\
&= E[X^2 - 2X\mu_X + \mu_X^2] \\
&= E[X^2] - 2E[X]\mu_X + \mu_X^2 \\
&= E[X^2] - \mu_X^2
\end{aligned}$$

(this latter formula is usually a more convenient way to find the variance.)

– Variance of X is 2nd central moment:

$$\begin{aligned}
\text{Var}[X] &= E[(X-\mu_X)^2] \\
&= E[X^2 - 2X\mu_X + \mu_X^2] \\
&= E[X^2] - 2E[X]\mu_X + \mu_X^2 \\
&= E[X^2] - \mu_X^2
\end{aligned}$$

(this latter formula is usually a more convenient way to find the variance.)

- The variance of a Gaussian random variable is the parameter σ^2 (you can get it through integration by parts or some clever manipulation)

PROPERTIES OF VARIANCE:

PROPERTIES OF VARIANCE:

1. $\text{Var}[c] = 0$

Properties of variance:

1. $\text{Var}[c] = 0$

2. $\text{Var}[X + c] = \text{Var}[X]$

PROPERTIES OF VARIANCE:

1. Var$[c] = 0$

2. Var$[X + c] =$ Var$[X]$

3. Var$[cX] = c^2$Var$[X]$

Hypothesis Tests Involving Sample Mean with Known Variance

Hypothesis Tests Involving Sample Mean with Known Variance

- Suppose we have two data sets that we think can be modeled as coming from Gaussian distributions

Hypothesis Tests Involving Sample Mean with Known Variance

- Suppose we have two data sets that we think can be modeled as coming from Gaussian distributions
- We compute the sample means by averaging the data sets

Hypothesis Tests Involving Sample Mean with Known Variance

- Suppose we have two data sets that we think can be modeled as coming from Gaussian distributions
- We compute the sample means by averaging the data sets
- If we observe a difference in the sample means for the two data sets, how can we determine analytically if the ensemble means are different?

- Let’s assume that distributions for the two data sets have a common variance **and we know the variance**, σ^2

- Let's assume that distributions for the two data sets have a common variance **and we know the variance**, σ^2

- We need to know a few more facts about sums of independent Gaussian random variables:

- Let's assume that distributions for the two data sets have a common variance **and we know the variance**, σ^2

- We need to know a few more facts about sums of independent Gaussian random variables:

1. If X and Y are independent RVs, then

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

2. If X and Y are independent Gaussian random variables with

$$X \sim \mathrm{Gaussian}\left(\mu_X, \sigma_X^2\right)$$

and

$$Y \sim \mathrm{Gaussian}\left(\mu_Y, \sigma_Y^2\right),$$

then

$$Z = X + Y \sim \mathrm{Gaussian}\left(\mu_Z, \sigma_Z^2\right).$$

2. If X and Y are independent Gaussian random variables with

$$X \sim \text{Gaussian}\left(\mu_X, \sigma_X^2\right)$$

and

$$Y \sim \text{Gaussian}\left(\mu_Y, \sigma_Y^2\right),$$

then

$$Z = X + Y \sim \text{Gaussian}\left(\mu_Z, \sigma_Z^2\right).$$

By linearity $\mu_z = \mu_X + \mu_Y$, and by the previous property $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$.

3. If Z is a random variable, $aZ + b$ is also a Gaussian random variable[1]

[1] All these things are proved in EEE 5544.

- Let's start by considering the statistics of a single sample mean,

$$\hat{\mu}_X = \frac{1}{m}\sum_{i=1}^{m} X_i$$

- By the properties above, we can see:

1. $\hat{\mu}_X$ is a Gaussian random variable

- Let's start by considering the statistics of a single sample mean,

$$\hat{\mu}_X = \frac{1}{m}\sum_{i=1}^{m} X_i$$

- By the properties above, we can see:

1. $\hat{\mu}_X$ is a Gaussian random variable
2. The mean of $\hat{\mu}_X$ is $E[\hat{\mu}_X] = \mu_X$

3. The variance of $\hat{\mu}_X$ is

$$\operatorname{Var}\left[\frac{1}{m}\sum_{\substack{0\\ i=}}^{m-1} X_i\right] = \frac{1}{m^2}\operatorname{Var}\left[\sum_{\substack{0\\ i=0}}^{m-1} X_i\right]$$

3. The variance of $\hat{\mu}_X$ is

$$\mathrm{Var}\left[\frac{1}{m}\sum_{0}^{m-1} X_i\right] = \frac{1}{m^2}\mathrm{Var}\left[\sum_{0}^{m-1} X_i\right]$$

$$= \frac{1}{m^2}\sum_{i=0}^{m-1}\mathrm{Var}\left[X_i\right]$$

by independ.

3. The variance of $\hat{\mu}_X$ is

$$
\begin{aligned}
\mathrm{Var}\left[\frac{1}{m}\sum_{0}^{m-1} X_i\right] &= \frac{1}{m^2}\mathrm{Var}\left[\sum_{0}^{m-1} X_i\right] \\
&= \frac{1}{m^2}\sum_{0}^{m-1}\mathrm{Var}\left[X_i\right] \\
&= \frac{1}{m^2}\sum_{0}^{m-1}\sigma_X^2
\end{aligned}
$$

3. The variance of $\hat{\mu}_X$ is

$$\begin{aligned}
\mathrm{Var}\left[\frac{1}{m}\sum_{0}^{m-1} X_i\right] &= \frac{1}{m^2}\mathrm{Var}\left[\sum_{0}^{m-1} X_i\right] \\
&= \frac{1}{m^2}\sum_{0}^{m-1}\mathrm{Var}\left[X_i\right] \\
&= \frac{1}{m^2}\sum_{0}^{m-1}\sigma_X^2 = \frac{\sigma_X^2}{m}
\end{aligned}$$

Variance of sample mean going down as $\frac{1}{m}$

(Note that the variance of the sample mean decreases linearly with the number of samples.

(Note that the variance of the sample mean decreases linearly with the number of samples. This can be used to show that the sample mean converges to the true mean if the variance of the original random variable is finite.)

EX

The city of Gainesville claims the mean commute time on SW 24th Ave from I-75 to UF is 23 minutes with a variance of 50. I traveled that route 10 times over the last two weeks and had an average commute time of 27 minutes. Conduct a hypothesis test to determine whether the City of Gainesville's model is reasonable. Reject the null hypothesis if $p < 0.01$.

Null hypothesis:

City's model is correct

$X_i \sim \text{Gaussian}(23, \sigma_X^2 = 50)$

What is the prob. that we observed a result this extreme, i.e. $\hat{\mu} \geq 27$ mins?

$$\hat{\mu} = \frac{1}{10} \sum_{i=1}^{10} X_i$$ (sample mean estimator

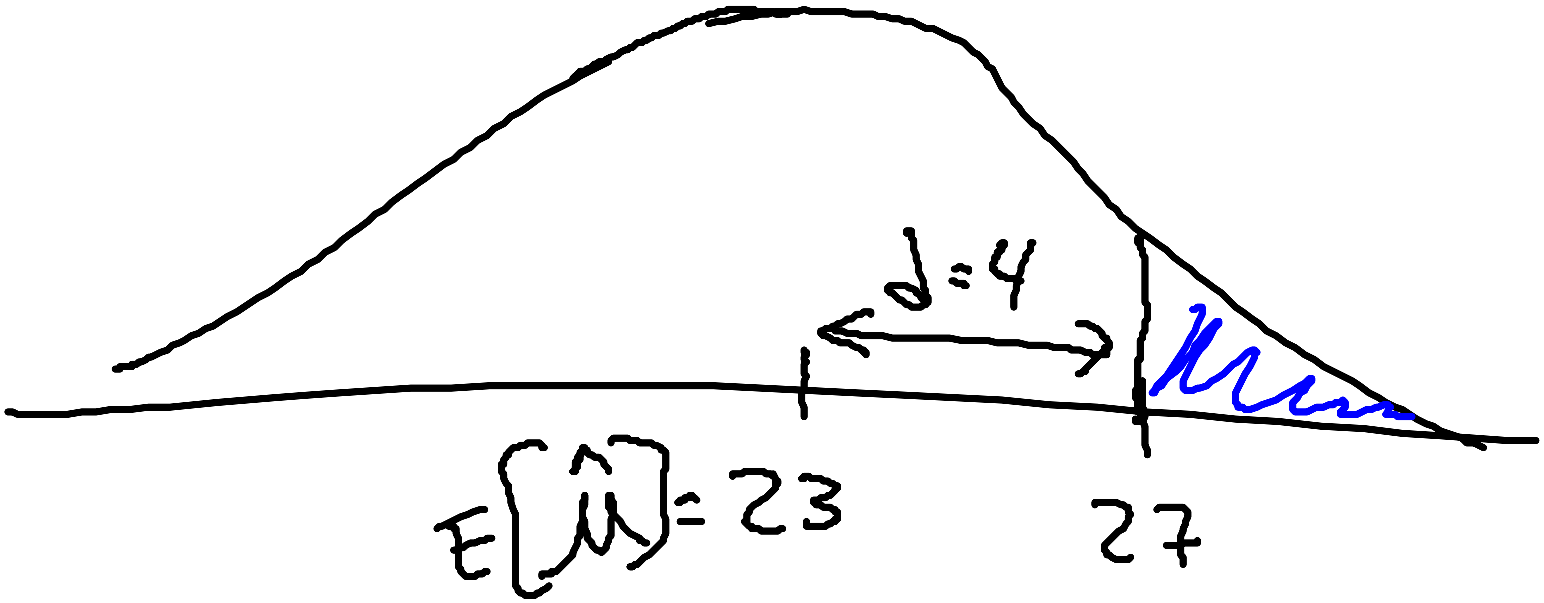
$$E[\hat{\mu}] = 23 = \mu$$

$$Var[\hat{\mu}] = \frac{\sigma^2}{10} = \frac{50}{10} = 5$$

$$P(\hat{\mu} \geq 27)$$

(one-sided hypothesis test

$$P(\hat{\mu} \geq 27)$$



$$p = P(\hat{\mu} \geq 27) = Q\left(\frac{d}{\sigma}\right) = Q\left(\frac{4}{\sqrt{5}}\right)$$