Lecture 8 (Ch. 2)

mean	variance	
$\bar{x}$	s^2	data / sample
		distr. / pop.

Keep the "big-picture" in mind ↗

We are looking at summary measures of data / sample.

Sample mean

$$\bar{x} = \frac{1}{n}\sum_{i}^{n} x_i$$

$\bar{x} \sim$ typical x

Sample variance

$$s^2 = \frac{1}{n-1}\sum_{i}^{n}(x_i - \bar{x})^2$$

$s \sim$ typical deviation in x

1 for-loop, instead of 2 for-loops.

Useful & fast formula for computing s^2:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2)$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - 2\bar{x}\underbrace{\sum_{i} x_i}_{n\bar{x}} + (\bar{x})^2 \overbrace{\sum_{i} 1}^{n}\right]$$

Step through this.

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - 2n(\bar{x})^2 + n(\bar{x})^2\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2\right]$$

$$= \frac{1}{n-1}\left[n\overbrace{\left(\frac{1}{n}\sum x_i^2\right)}^{\overline{x^2}} - n(\bar{x})^2\right] = \frac{n}{n-1}\left[\overline{(x^2)} - (\bar{x})^2\right]$$

$$s^2 = \frac{1}{n-1}\sum_{i}(x_i - \bar{x})^2 \quad \text{"Defining formula"}$$

$$s^2 = \frac{n}{n-1}\left[\overline{x^2} - (\bar{x})^2\right] \quad \text{"Computational formula"}$$

$\boxed{\text{Example}}$

$x = c(1, 3, 8) \rightarrow x^2 = c(1, 9, 64) \rightarrow \overline{x^2} = \frac{74}{3}$

$$s^2 = \frac{3}{2}\left[\frac{74}{3} - 16\right] = \frac{3}{2}\,\frac{74-48}{3} = \frac{26}{2} = 13$$

Now, we need to come-up with the corresponding Things in The $\boxed{\text{pop.}}$

So, switch to distributions ($p(x)$, $f(x)$). No Data!

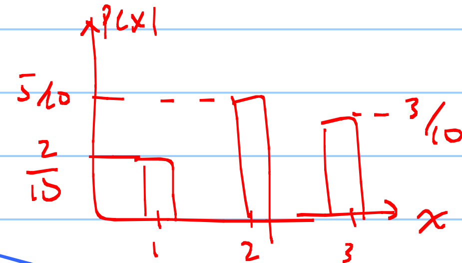
1) Distribution mean (or Expected Value) $= E[x] = \mu_x = \begin{cases} \sum_x x\, p(x) \\ \int x\, f(x)\, dx \end{cases}$

Motivation: Consider a "pop." of size 10:

$\{3, 2, 2, 1, 3, 2, 3, 1, 2, 2\}$

$$\text{mean} = \frac{1}{10}[3+2+2+\cdots] = \frac{1}{10}[3(3) + 5(2) + 2(1)]$$

$$= \frac{3}{10}(3) + \frac{5}{10}(2) + \frac{2}{10}(1) = \sum_{\forall x} p(x)\cdot x \quad , \text{where}$$



Compare:

Sample mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$,

distr. mean (Expected Value): $E[x] \equiv \mu_x = \sum_x x\, p(x)$, $\int_{-\infty}^{\infty} x\, f(x)\, dx$

the book drops the x on μ_x, but then μ can be confused with the parameter of the Normal distr.

FYI

$E[x]$ does not mean that E is a function of x. In fact, E is a $\sum_x$ or an $\int_{-\infty}^{\infty} dx$, and so it is not a function of x. $E[x]$ simply means that you need $p(x)$ or $f(x)$ to find it.

See binomial example, below.

Example Binomial (n, π)

$$E[x] = \sum_{x=0}^{n} \frac{n!}{x!(n-x)!} \pi^x (1-\pi)^{n-x} \cdot x$$

or μ_x

x=0 contributes zero to the sum

relabel $\sum_x$ and note that $\frac{x}{x!} = \frac{1}{(x-1)!}$

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} \pi^x (1-\pi)^{n-x}$$

$y = x-1$

$$= \sum_{y=0}^{n-1} \frac{(n-1)!\, n}{y!(n-y-1)!} \pi^{y+1} (1-\pi)^{n-y-1}$$

$$= \underline{n\pi} \sum_{y=0}^{n-1} \frac{(n-1)!}{y!(n-y-1)!} \pi^y (1-\pi)^{n-y-1}$$

$m = n-1$

$$= (m+1)\pi \sum_{y=0}^{m} \frac{m!}{y!(m-y)!} \pi^y (1-\pi)^{m-y}$$

$$= \underbrace{(m+1)}_{n} \pi \underbrace{\sum_{y=0}^{m} \binom{m}{y} \pi^y (1-\pi)^{m-y}}_{=1=\sum_{y=0}^{n} p(y)}$$

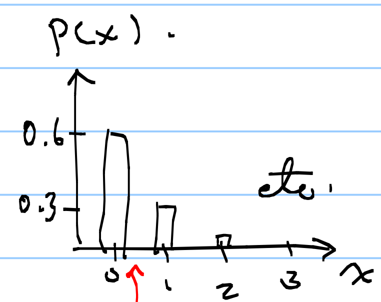
$$\boxed{E[X] = n \cdot \pi}$$

2 params of binomial Note

Note $E[x]$ is not a function of x.

E.g. 1.23:

# of Bads out of 100 → x	0	1	2	3	4
$p(x)$	.6058	.3044	.0757	.0124	---



Hard way.

$$E[x] = \sum x\, p(x) = 0(.6058) + 1(.3044) + \ldots$$

$$= n\pi = 100(.005) = 0.5$$

Easy way.

On avg. 0.5 out of 100 (i.e. 1 out of 200) computers are defective.

For the other distributions, same tricks:

Poisson(λ) : $\mu_x = E[x] = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \cdots = \lambda \underbrace{\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}}_{1 = \sum_x p(x)} = \lambda$

Now you can see why λ is called mean.

Normal(μ, σ):

change of variables.

$$\mu_x = E[x] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \cdots \left(z = \frac{x-\mu}{\sigma}\right) \cdots = \mu \cdot 1$$

$\int f(x)\,dx = 1$

Now you can see why μ (The param of Normal) is a mean.

Etc. We can find The mean of any distribution in terms of parameters of That distr.

Warning: Don't confuse $\bar{x}$, μ_x, μ

Sample mean. $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

distr. mean

$E[x] = \mu_x = n\pi$ binomial (n, π)

$\mu_x = \lambda$ poisson(λ)

$\mu_x = \mu$ Normal(μ, σ)

Note about $\sum_x x\, p(x)$:

Recall That $p(x)$ is The mass function, where x = discrete/Categ.

E.g. x = Computer brand = {Apple, Dell, Lenovo}

or x = Speed = {100, 200, 300} miles per hour.

quantitative qualitative. (see lect 1).

$\sum_x x\, p(x)$ makes sense only for x = quantitative (e.g. binomial)

Because of my typo, This qz will not be counted at all.

$\boxed{\text{Q1:}}$ Let x be a r.v. taking values 0, 1, 2. We have observed x five times and have found The values 0, 1, 0, 1, 2; we know That The $p(x)$ is $p(x=0)=\frac{2}{3}$, $p(x=1)=\frac{1}{6}$, $p(x=2)=\frac{1}{6}$.

The distr. mean of x is

$\boxed{\boxed{\text{A) } \cancel{1/3}}\ 1/2}$ B) 4/5 C) 1 D) Does not exist.

$$\mu_x=\sum_{x=0}^{2} x\,p(x) = 0\,p(0)+1\,p(1)+2\,p(2) = \frac{1}{6}+\frac{2}{6} = \frac{1}{2}$$

$$\bar{x}=\frac{1}{5}(0+1+0+1+2)=4/5 \qquad \frac{1}{3}(0+1+2) = 1?$$

Summary

Single summary of histogram location

Sample mean:

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i$$

~ typical x/obs.

Single summary of histogram spread

Sample variance:

recall computational formula, too. →

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Sample std. dev. = s.

~ typical deviation/spread

Single summary of distribution/population location

dist./pop mean, or $E[x]$

$$\mu_x = E[x] = \sum_x x\,p(x), \quad \int_{-\infty}^{\infty} x f(x)\,dx$$

Single summary of distr./pop. spread

dist./pop. Variance

$$\sigma_x^2 = V[x] = \begin{cases} \sum_{\forall x}(x-E[x])^2 p(x) \\ \int_{-\infty}^{\infty}(x-E[x])^2 f(x)\,dx \end{cases}$$

$E[x]=\sum x\,p(x)$

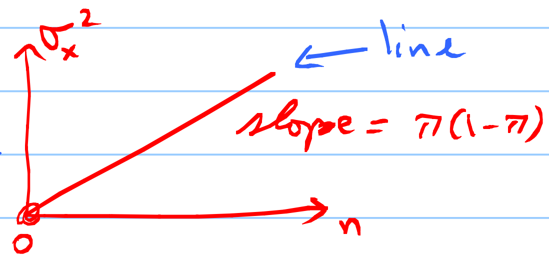
New

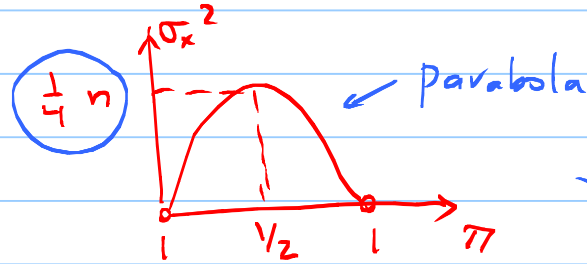
Don't drop This x, like The book does.

Similarly to sample std. dev., The pop. std. dev. is σ_x.

Let's find The $\sigma_x^2 = V[x]$ of our special dists:

Binomial(n,π) : $V[x] \equiv \sigma_x^2 = \sum_x (x - \underbrace{n\pi}_{\mu_x \text{ of Binomial}})^2 \underbrace{p(x)}_{p(x) \text{ of Bin.}} = \cdots = \underline{n\pi(1-\pi)}$.





Interpretation: σ_x is The (expected) typical deviation in x.

So, $\sigma_x \sim \sqrt{n}$ The maximum σ_x is $\frac{1}{2}\sqrt{n}$

$\therefore$ If you are tossing n coins, The Typical x (# of heads) is about $n\pi$ (ie. μ_x), and The typical dev. in x is at most $\frac{1}{2}\sqrt{n}$.

Poisson(λ): $\sigma_x^2 = V[x] = \sum (x-\lambda)^2 \underbrace{p(x)}_{\text{poisson}} = \cdots = \lambda$

Recall $E[x] = \lambda$ ⟵ same ⟶

$E[x] = \mu$

Normal(μ,σ): $\sigma_x^2 = V[x] = \int (x-\mu)^2 \underbrace{f(x)}_{\text{Normal}} dx = \cdots = \sigma^2$

which is why the param. σ^2 is called variance.

Example: $f(x) = 2x$, $0 < x < 1$. $\mu_x = \int x f(x)\,dx = \cdots = \frac{2}{3}$.

Find $V[x] \equiv \sigma_x^2$

A) 1   B) $\frac{1}{2}$ $\boxed{\text{C) } \frac{1}{18}}$ D) This $f(x)$ has no variance.

$$V[x] \equiv \sigma_x^2 = \int_{-\infty}^{\infty} (x-\mu_x)^2 f(x)\,dx = \int_0^1 \left(x - \frac{2}{3}\right)^2 2x\,dx = 2\int_0^1 \left(x^3 - \frac{4}{3}x^2 + \frac{4}{9}x\right) dx$$

$$= 2\left[\frac{1}{4}x^4 - \frac{4}{3}\frac{1}{3}x^3 + \frac{4}{9}\frac{1}{2}x^2\right]_0^1 = 2\left[\frac{1}{4} - \frac{4}{9} + \frac{2}{9}\right] = 2\left(\frac{1}{4} - \frac{2}{9}\right) = \frac{2}{36} = \frac{1}{18}$$

By now, you should be familiar with the meaning of

histograms vs. distributions

Sample mean $\bar{x}$ vs. distr. mean $E[x] = \mu_x$

" Variance s^2 vs. " Variance $V[x] = \sigma_x^2$

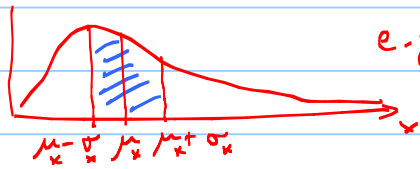
" std. dev. s vs. " std. dev. σ_x

Finally, given that we can compute all of the above quantities, you can then compute the proportion of times x is expected to be within some std. dev. of its mean, for ANY distr.

1, 1.96, 2, ...

For examples, for the normal dist. we can now say that 68% of x's fall within 1 std. dev. of the mean.

But now we can say things like that for any distr. even skewed ones:



e.g. Poisson, ...

Computing areas like this will eventually enable us to provide some measure of confidence when we try to estimate a population parameter, later.

We use the sample mean ($\bar{x}$) to estimate the pop. mean (μ_x), and the sample variance (s^2) to estimate the pop. variance (σ_x^2). More, later.

Summary

Single summary of histogram location

Sample mean:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$\sim$ typical x/obs.

recall computational formula, too. →

Single summary of histogram spread

Sample variance:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Sample std. dev. = s.

$\sim$ typical deviation/spread

Recall why $\frac{1}{n}\sum(x_i-\bar{x})$ will not do!

Single summary of distribution/population location

dist./pop mean, or $E[X]$

$$\mu_x \equiv E[X] = \sum_x x\, p(x) \quad . \quad \int_{-\infty}^{\infty} x f(x)\, dx$$

Single summary of dist/pop. spread

dist/pop. var. or $V[X]$

$$\sigma_x^2 \equiv V[X] = \sum_x (x-\mu_x)^2 p(x)$$

$$\int_{-\infty}^{\infty} (x-\mu_x)^2 f(x)\, dx$$

Eg. binomial (n, π): $\mu_x = n\pi$ $\qquad \sigma_x^2 = n\pi(1-\pi)$

poisson (λ): $\mu_x = \lambda$ $\qquad \sigma_x^2 = \lambda$

Normal (μ, σ): $\mu_x = \mu$ $\qquad \sigma_x^2 = \sigma^2$

uniform (a, b): $\mu_x = \frac{a+b}{2}$ $\qquad \sigma_x^2 = \frac{(b-a)^2}{12}$

Exponential (λ): $\mu_x = \frac{1}{\lambda}$ $\qquad \sigma_x^2 = \left(\frac{1}{\lambda}\right)^2$

hw-lect 8-1

Consider the binomial distr. $p(x)$ with parameters $n=4$, $\pi=\frac{1}{4}$.

a) Compute specific values of $p(x)$ for all possible values of x. (By hand or By R).

b) Compute $E[x] \equiv \sum_x x\, p(x)$, and compare the answer with the value of $(n\pi)$. (By hand or By R).

c) Take a sample of size 100 from $p(x)$, compute the sample mean of the 100 numbers, and compare the answer with the answer in part b. (By R)

hw-lect 8-2

For the uniform distr. (see 1.19) between a, b, show that the expected value is $\frac{1}{2}(a+b)$, and the variance is $\frac{1}{12}(b-a)^2$.

hw-lect 8-3

For the exponential distr. with param. λ, find μ_x and σ_x^2.

Hints: $\int_0^\infty y e^{-y}\, dy = 1$ $\int_0^\infty (y-1)^2 e^{-y}\, dy = 1$

Don't do This

Find the μ_x (not σ_x, it's too long!) for

a) The $p(x)$ given in exercise 1.27, with the two "?" given as 0.1, and zero, respectively.

b) The $f(x)$ given in exercise 1.21

Don't do This

This exercise will help to get a better sense of what σ_x measures, geometrically.

Consider $f(x) = \begin{cases} 1+x & -1 < x \le 0 \\ 1-x & 0 < x < 1 \\ 0 & \text{else} \end{cases}$

a) Plot/Graph $f(x)$ vs. x

b) Confirm That $f(x)$ is a density function.

c) Compute The mean μ_x.

d) compute the variance σ_x^2.

Don't forget to check The soln (when it's posted) to see an interpretation of σ_x.

hw-lect8-4) Find The area within one standard deviation of The mean (ie. $\mu_x \pm \sigma_x$) for

a) binomial ($n=20$, $\pi = \frac{1}{4}$)
b) poisson ($\lambda = 5$)
c) Normal($\mu=5$, $\sigma=1$)

hw-lect8-5

In Example 1.23 (in Text and in Lect), we found That on the average, out of 100 computers, 0.5 computers are defective.

a) What is The typical deviation we expect to see from This number (still out of 100)?

b) Suppose we do not know That The proportion of defective computers is 0.005. Then out of 100 computers, what is The maximum value we expect to see for typical deviation?