Lecture 10 (Ch-2)

Last time: $\bar{x} \sim$ typical x (sample mean of x), $s \sim$ typical deviation (sample std. dev. of x).

Now, we need to come-up with the corresponding things in the pop.

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

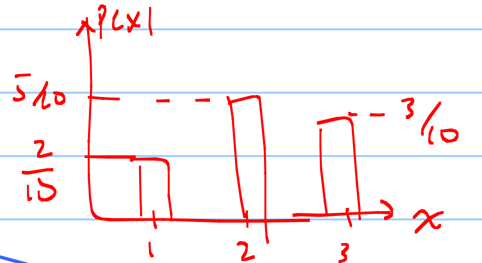
1) Expected Value (or mean) $= E[x] = \mu_x = \begin{cases} \sum_x x\, p(x) \\ \int x\, f(x)\, dx \end{cases}$

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 with Distribution mean, or Expected Value

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

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

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

Note: $E[x]$ 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 $\frac{x}{x!} = \frac{1}{(x-1)!}$

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

$y = x-1$

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

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

$m = n-1$

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

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

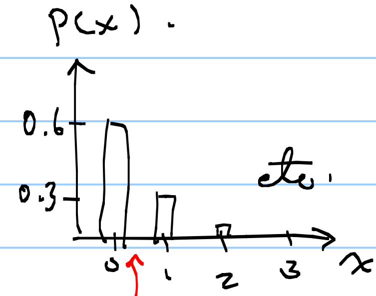
$$\boxed{E[X] = n \cdot \pi}$$

2 params of binomial Note

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

E.g. 1.22:

# 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.

$\boxed{Q1:}$ Let $f(x)=2x$, $0<x<1$. $E[x]$ or μ_x is

A) 1 $\quad$ B) $\frac{2}{3}$ $\quad$ C) $\frac{1}{2}$ $\quad$ D) This $f(x)$ has no mean.

$$E[x] = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_0^1 x\, 2x\, dx = 2\,\frac{1}{3}x^3 \Big|_0^1 = \frac{2}{3}(1-0)$$

For the other distributions, same tricks:

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

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

$E[x] = \mu_x = n\pi$ $\quad$ binomial (n, π)

$\mu_x = \lambda$ $\quad$ poisson(λ)

$\mu_x = \mu$ $\quad$ 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 = fruit type = {Apple, Orange, Kiwi}

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)

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 $E[X]$

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

Eg. binomial (n, π): $\mu_x = n\pi$

poisson (λ): $\mu_x = \lambda$

Normal (μ, σ): $\mu_x = \mu$

uniform (a,b): $\mu_x = \frac{a+b}{2}$

Exponential (λ): $\mu_x = \frac{1}{\lambda}$

hw

Single summary of distr./pop. spread

Next time!

hw_lect10_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 $E[x] = \sum_{x} x\, p(x)$, 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_lect10_2

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

hw_lect10_3

Find the expected value of the exponential distr. with param. λ

Hint: $\int_0^\infty y e^{-y}\, dy = 1$

hw_lect10-4

Find the μ_x 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 1.22.

This document was created with Win2PDF available at http://www.win2pdf.com.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.
This page will not be added after purchasing Win2PDF.