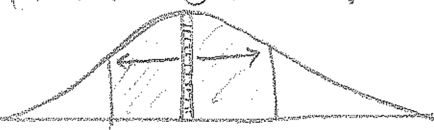
DAY 26
WED NOV 4

$\mathbb{E}$XPECTATION & VARIANCE



We already "know" what expectation and variance are, but in a slightly different context. But let's understand them in terms of probability first.

If $\rho(x)$ is a P.D.F. on $[a,b]$, and x is chosen randomly" via ρ, then the expectation $\mathbb{E}[f(x)]$ is the AVERAGE VALUE of $f(x)$

$$\boxed{\mathbb{E}[f(x)] = \int_a^b f(x) \cdot \rho(x)\, dx}$$

Same as the "mass" of $f(x)$ given density $\rho(x)$.

for $f(x) = x$, this is called the MEAN, $\bar{x}$.

Q1. Given PDF $\rho(x) = \sin(x)$ on $[0, \pi/2]$, find the average of $\cos(x)$ for x chosen randomly according to ρ.

Ans

$$\mathbb{E}[\cos(x)] = \int_0^{\pi/2} \cos(x) \cdot \sin(x)\, dx$$

[see the formula above]

$$= \tfrac{1}{2} \int_0^{\pi/2} \sin(2x)\, dx$$

$$= -\tfrac{1}{4} \cos(2x) \Big|_{x=0}^{x=\pi/2}$$

$$= \boxed{1/2}$$

WARNING

$\mathbb{E}[f(x)]$ is <u>NOT</u> necessarily the same as $f(\mathbb{E}[x])$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = f(\bar{x})$

In the example above,

$$\bar{x} = \int_0^{\pi/2} x \sin x\, dx$$

(by parts)

$$= \left[-x\cos(x)\right]_0^{\pi/2} + \int_0^{\pi/2} \cos(x)\,dx$$

$$= 0 + \left[\sin x\right]_0^{\pi/2} = 1.$$

So, $f(\bar{x}) = \cos(1) \neq 1/2.$

MORAL:

Average of function is <u>NOT</u> function of the average

(More on this later)

Q2 | Consider $p(x) = c/x^3$ on $[1,\infty]$. Find the c which makes this a PDF, and compute $\overline{x}$

Ans. $$\int_1^{\infty} x^{-3}\,dx = \left[\frac{x^{-3+1}}{-3+1}\right]_{x=1}^{x=\infty} = \tfrac{1}{2}\left[0-(-1)\right] = \underline{\tfrac{1}{2}}$$

So, $c=2$, and we only need to compute $\overline{x}$.

$$\overline{x} = \mathbb{E}[x] = \int_1^{\infty} x \cdot \left(2x^{-3}\,dx\right)$$
$$= 2\int_1^{\infty} x^{-2}\,dx$$
$$= 2\cdot\left[\frac{x^{-2+1}}{-2+1}\right]_{x=1}^{x=\infty} = 2\left(0-(-1)\right) = \boxed{2}$$

Make sure you can also compute $\mathbb{E}[x \sin x]$, etc. above.

Q3 | Repeat Q2, but $p(x) = c\cdot x(1-x)$ on $[0,1]$

Ans. $$\int_0^1 x(1-x)\,dx = \int_0^1 (x-x^2)\,dx = \left(x^2/2 - x^3/3\right)_{x=0}^{x=1}$$
$$= \tfrac{1}{2} - \tfrac{1}{3} = \underline{\tfrac{1}{6}} \qquad \text{So, } \underline{c=6}.$$

Now, $$\overline{x} = \int_0^1 x\cdot\left[6x(1-x)\right]dx$$
$$= 6\int_0^1 (x^2-x^3)\,dx = 6\left[x^3/3 - x^4/4\right]_{x=0}^{x=1}$$
$$= 6\left(\tfrac{1}{3}-\tfrac{1}{4}\right) = 6/12 = \boxed{1/2}$$

VARIANCE: What is the "AVERAGE SQUARED DISTANCE" between x and $\overline{x}$?

$$V(x) = \mathbb{E}\left[(x-\overline{x})^2\right] = \mathbb{E}\left[x^2 - 2x\overline{x} + \overline{x}^2\right]$$

$$= \int_a^b (x^2 - 2x\bar{x} + \bar{x}^2)\,\rho(x)\,dx$$

$$= \int_a^b x^2 \rho(x)\,dx - 2\bar{x}\underbrace{\int_a^b x\rho(x)\,dx}_{\mathbb{E}[x]=\bar{x}} + \bar{x}^2 \underbrace{\int_a^b \rho(x)\,dx}_{1,\ \rho \text{ is a PDF}}$$

$$= \boxed{\mathbb{E}[x^2] - \left(\mathbb{E}[x]\right)^2}$$

More generally,

$$\mathbb{V}[f(x)] = \mathbb{E}[f^2(x)] - \left(\mathbb{E}[f(x)]\right)^2$$

Q4. What is the variance of $\cos(x)$ in Q1?

Ans

$$\mathbb{V}[\cos(x)] = \mathbb{E}[\cos^2(x)] - \left(\mathbb{E}[\cos(x)]\right)^2 \longleftarrow \text{know this from Q1}$$

$$= \int_0^{\pi/2} \cos^2(x) \cdot \sin(x)\,dx - 1/4$$

$= $ (easy)

Q5. Find $\mathbb{V}[x]$ in Q2.

Ans.

$$\mathbb{V}[x] = \mathbb{E}[x^2] - \left(\mathbb{E}[x]\right)^2 \quad \text{(from Q2)}$$

$$= \int_1^\infty x^2 \cdot x^{-3}\,dx - 4$$

$$= \int_1^\infty x^{-1}\,dx - 4 = \underline{\underline{\infty}}$$

On average, x is INFINITELY far from the average!!

Q6. Find $\mathbb{V}[x]$ in Q3!

Ans

$$\mathbb{V}[x] = \mathbb{E}[x^2] - \left(\mathbb{E}[x]\right)^2 = \int_0^1 x^2 \cdot (6x(1-x))\,dx - 1/4$$

$$= 6 \int_0^1 (x^3 - x^4)\,dx - 1/4$$

$$= 6 \cdot \left(x^4/4 - x^5/5 \right)_{x=0}^{x=1} - 1/4$$

$$= 6 \cdot (1/4 - 1/5) - 1/4 = 6/20 - 1/4 = \boxed{\frac{1}{20}}$$

FRESHMAN USES "ONE WEIRD POLAR TRICK". PENN PROFESSSORS HATE HER! (Bonus!!)

The GAUSSIAN or NORMAL distribution is EVERYWHERE, it has PDF $\rho(x) = ce^{-x^2}$ on $(-\infty, \infty)$, but we can't figure out "c" without this kool use of polar coordinates:

$$I = \int_{-\infty}^{\infty} e^{-x^2}\,dx$$

so, $$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2}\,dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2}\,dy \right)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,dx\,dy \quad \leftarrow \text{"double" integral! over the plane.}$$

Polar coordinates! $r = x^2 + y^2$ (0 to ∞)
$\theta = \arctan(y/x)$ (0 to 2π)

$\leftarrow$ $dx\,dy$ becomes $\underline{r}\,d\theta\,dr$.

d Area = $r\,dr\,d\theta$.

So, $$I^2 = \int_0^{\infty} \int_0^{2\pi} e^{-r^2} \cdot r\,d\theta\,dr$$

(integrate $\int_0^{2\pi} d\theta$ out)

$$= 2\pi \int_0^{\infty} e^{-r^2} \cdot r\,dr$$

u-sub! $u = r^2$, $du = \frac{1}{2} r\,dr$, get

$I^2 = \pi$, so $I = \sqrt{\pi}$.

So, GAUSSIAN: $$\boxed{\rho(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}}$$