§3.3 Expected Values (and Variance)

Def: The expected value of a discrete r.v. X is

$$E[X] = \sum_x x \cdot P(X = x)$$

The expected value of a function $H(X)$ is similar

$$E[H(X)] = \sum_x x \cdot P(H(X) = x)$$

$$= \sum_x H(x) \cdot P(X = x)$$

For example $E[X^2] = \sum_x x^2 \cdot P(X = x)$

Note: For $E[X]$ we often write μ_X or just μ and say "mean" or "average"

↳ $E[X]$ is a measure of the "center" of the distrib.

Why call $E[X]$ the mean?

→ Suppose X has possible outcomes 1, 2, 3

with
$$P(X=1) = f(1) = 1/10$$
$$P(X=2) = f(2) = 3/10$$
$$P(X=3) = f(3) = 6/10$$

Then sampling X 10 times should yield

$$\left.\begin{array}{l} 1 \times 1 \text{ time} \\ 2 \times 3 \text{ times} \\ 3 \times 6 \text{ times} \end{array}\right\} \text{Average Value} = \frac{1\times 1 + 2\times 3 + 3\times 6}{10}$$

$$= 1\cdot 1/10 + 2 \cdot 3/10 + 3 \cdot 6/10$$

Note: $E[X]$ vs. $E[H(X)]$ is usually complicated...

but $E[aX + b] = a\,E[X] + b$

(expected value is linear)

↳ More generally $E[X + Y] = E[X] + E[Y]$

Example: Household size in Turkey is distributed as

Size	1	2	3	4	5	6	7	8+
Prob	8.5%	22.3%	21.0%	23.7%	12.4%	5.5%	28%	3.8

Mean household size is

$$E[X] = 1(.085) + 2(.223) + 3(.21) + 4(.237)$$
$$+ 5(.124) + 6(.055) + 7(.028) + 8(.038)$$
$$= 3.56$$

Example: Roll die. X = #rolled

$$E[X] = 1\cdot\tfrac{1}{6} + 2\cdot\tfrac{1}{6} + 3\cdot\tfrac{1}{6} + 4\cdot\tfrac{1}{6} + 5\cdot\tfrac{1}{6} + 6\cdot\tfrac{1}{6}$$
$$= \frac{21}{6} = \frac{7}{2} = 3.5$$

Example: Flip coins until H. X = #flips

$$E[X] = 1\cdot\tfrac{1}{2} + 2\cdot\tfrac{1}{4} + 3\cdot\tfrac{1}{8} + 4\cdot\tfrac{1}{16} + \cdots$$

$= \cdots$ (tricky math or use computer) $\cdots = 2$

Example: (St. Petersburg Paradox)

Gamble on coin flipping game.

- Pay 5 TL to play
- Flip coin until H
- Win $2^{\#\text{flips}-1}$ TL

Ex: Get H in 1 flip $\Rightarrow$ win 1 TL
2 flips $\Rightarrow$ win 2 TL
3 flips $\Rightarrow$ win 4 TL
$\vdots$

Let X = money won.

$$E[X] = 1\cdot\tfrac{1}{2} + 2\cdot\tfrac{1}{4} + 4\cdot\tfrac{1}{8} + 8\cdot\tfrac{1}{16} + \cdots$$
$$= \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \cdots$$
$$= \infty \quad !!$$

$$E[\text{total income}] = E[X-5] = \infty - 5 = \infty \quad !!$$

Note: To actually get near this "expected value", you need to play the game infinitely many times. If you only have a finite amount of money then you will likely lose it all in this game...

Def: The variance of a discrete r.v. X is

$$\mathrm{Var}[X] = E[(X-\mu)^2] = \sum_x (x-\mu)^2 \cdot P(X=x)$$
$$\underset{\text{Thm.}}{=} E[X^2] - (E[X])^2$$

The standard deviation is $\sigma_X = \sqrt{\mathrm{Var}[X]}$

Note: We often write σ_X^2 or just σ^2 for $\mathrm{Var}[X]$

$\hookrightarrow$ $\mathrm{Var}[X]$ is a measure of how "spread out" the distribution is. (Expected dist^2 from mean)

As with $E[X]$, $\mathrm{Var}[H(X)]$ is usually messy...

but $\mathrm{Var}[aX+b] = a^2\,\mathrm{Var}[X]$

(i.e. $\sigma_{aX+b}^2 = a^2\sigma_X^2$

$\sigma_{aX+b} = |a|\cdot\sigma_X$)

It is usually very annoying to compute σ, but is sometimes useful for quickly solving a problem

$\rightarrow$ see "68-95-99.7 Rule"

(1, 2, 3 standard dev. of normal)

$\rightarrow$ see Chebyshev's Inequality $P(|x-\mu| \geq n\sigma) \leq 1/n^2$