Probability review II: Random variables, expectation, and variance

Topics we'll cover

1. What is a random variable?
2. Expected value
3. Variance and standard deviation

Random variables

Roll two dice. Let X be their sum.

$$\begin{aligned} \text{outcome} = (1,1) \quad &\Rightarrow \quad X = 2 \\ \text{outcome} = (1,2) \text{ or } (2,1) \quad &\Rightarrow \quad X = 3 \end{aligned}$$

Probability space:

- Sample space: $\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$.
- Each outcome equally likely.

Random variable X lies in $\{2,3,4,5,6,7,8,9,10,11,12\}$.

A **random variable (r.v.)** is a defined on a probability space.
It is a mapping from Ω (outcomes) to $\mathbb{R}$ (numbers).
We'll use capital letters for r.v.'s.

The distribution of a random variable

Roll a die.

Define $X = 1$ if die is ≥ 3, otherwise $X = 0$.

$X \in \{0,1\}$

$P(X=0) = P(\text{die} = 1,2) = 1/3$

$P(X=1) = P(\text{die} = 3,4,5,6) = 2/3$

Expected value, or mean

Expected value of a random variable X:

$$\mathbb{E}(X) = \sum_{x} x \underbrace{\Pr(X = x)}.$$

p

Roll a die. Let X be the number observed. $X \in \{1, 2, 3, 4, 5, 6\}$
What is $\mathbb{E}(X)$?

$$E(x) = 1\cdot\tfrac{1}{2} + 2\cdot\tfrac{1}{6} + 3\tfrac{1}{6} + 4\tfrac{1}{6} + 5\tfrac{1}{6} + 6\tfrac{1}{6} = \frac{21}{6} = 3.5$$

$$E(x) = 3.5$$

Another example

A biased coin has heads probability p.
Let X be 1 if heads, 0 if tails. What is $\mathbb{E}(X)$?

$$\begin{aligned} E(X) &= 0 \cdot P(X=0) + 1 \cdot P(X=1) \\ &= 0 \cdot (1-p) + 1 \cdot p \end{aligned}$$

$$E(X) = p$$

A property of expected values

How is the average of a set of numbers affected if:

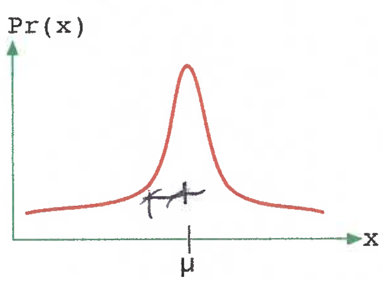
- You double the numbers?
- You increase each number by 1?

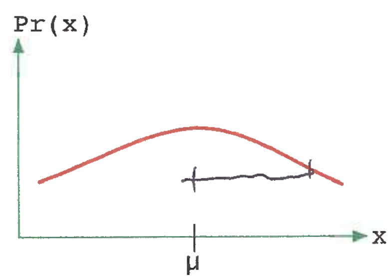
Summary: Let X be any random variable.
If $V = aX + b$ (any constants a, b), then $\mathbb{E}(V) = a\mathbb{E}(X) + b$

$E(V) = a\,E(X) + b$

Variance

Can summarize an r.v. X by its mean, μ. But this doesn't capture the **spread** of X:





A measure of spread: average distance from the mean, $\mathbb{E}(|X - \mu|)$? Mathematically, it is more convenient to look at average square distance from the mean - VARIANCE

- **Variance:** $\boxed{\mathrm{var}(X) = \mathbb{E}((X-\mu)^2),}$ where $\mu = \mathbb{E}(X)$
- **Standard deviation** $\sqrt{\mathrm{var}(X)}$:
  Roughly, the average amount by which X differs from its mean.

Variance: example

Choose X uniformly at random from $\{1, 2, 3, 4, 5\}$.

$E(X) = 1 \cdot \frac{1}{5} + 2\,{}^{1}/_{5} + 3\,{}^{1}/_{5} + 4\,{}^{1}/_{5} + 5\,{}^{1}/_{5} = 3 = \mu$

$Var = E\left[(X-\mu)^2\right] = 0 \cdot \frac{1}{5} + 1 \cdot {}^{2}/_{5} + 4 \cdot \frac{2}{5} = 2$

X	$(X-\mu)^2$	
1	2^2	$(1-3)^2 = (-2)^2$
2	1^2	$(2-3)^2 = (-1)^2$
3	0^2	$(3-3)^2 = 0^2$
4	1^2	$(4-3)^2 = 1^2$
5	2^2	$(5-3)^2 = 2^2$

$Var(X) = 2$

Std Deviation $= \sqrt{Var(X)} = \sqrt{2}$

$\mu = 3$

$Std(X) = \sqrt{2}$

$Var(X) = 0 \cdot \frac{1}{5} + 1\left(\frac{1}{5}\right) + 1\left(\frac{1}{5}\right) + 4\left(\frac{1}{5}\right) + 4\left(\frac{1}{5}\right) =$

$= 0 \cdot \frac{1}{5} + 1\left(\frac{2}{5}\right) + 4\left(\frac{2}{5}\right) = 2$

Variance: properties

Variance: $\text{var}(X) = \mathbb{E}((X-\mu)^2)$, where $\mu = \mathbb{E}(X)$

- Variance is always ≥ 0
- How is the variance affected if: } Properties
  - You increase each number by 1? ← not affected
  - You double each number? ← affected
- Summary: If $V = aX + b$ then $\text{var}(V) = a^2\,\text{var}(X)$

Std Dev is multiplied by a^2 if we multiply the variance by a

Alternative formula for variance

Variance: $\text{var}(X) = \mathbb{E}((X-\mu)^2)$, where $\mu = \mathbb{E}(X)$

Another way to write it: $\text{var}(X) = \mathbb{E}(X^2) - \mu^2$

Example: Choose X uniformly at random from $\{1, 2, 3, 4, 5\}$.

$$E(X^2) = 1^2 \cdot \frac{1}{5} + 2^2 \cdot \frac{1}{5} + 3^2 \cdot \frac{1}{5} + 4^2 \cdot \frac{1}{5} + 5^2 \cdot \frac{1}{5} =$$

$$= \frac{1+4+9+16+25}{5} = 11$$

$$Var(X) = E(X^2) - \mu^2 = 11 - 3^2 = 11 - 9 = 2$$

$$Var(X) = 2$$