Probability

3.1 Discrete Random Variables Basics

Anna Karlin
Most slides by Alex Tsun

Agenda

- Recap on rvs
- Expectation
- Linearity of Expectation (LoE)
- Law of the Unconscious Statistician (Lotus)

Random Variable

Suppose we conduct an experiment with sample space Ω. A **random variable (rv)** is a numeric function of the outcome, $X: \Omega \to \mathbb{R}$. That is, it maps outcomes $\omega \in \Omega$ to numbers, $\omega \mapsto X(\omega)$.

The set of possible values X can take on is its **range/support**, denoted Ω_X.

If Ω_X is finite or countably infinite (typically integers or a subset), X is a **discrete** random variable (drv). Else if Ω_X is uncountably large (the size of real numbers), X is **continuous** random variable.

Probability Mass Function (PMF)

The **probability mass function (pmf)** of a discrete random variable X assigns probabilities to the possible values of the random variable. That is, $p_X: \Omega_X \to [0,1]$ where

$$p_X(k) = P(X = k)$$

Note that $\{X = a\}$ for $a \in \Omega_X$ form a partition of Ω, since each outcome $\omega \in \Omega$ is mapped to exactly one number. Hence,

$$\sum_{z \in \Omega_X} p_X(z) = 1$$

Homeworks of 3 students returned randomly

- Each permutation equally likely
- X: # people who get their own homework

Prob	Outcome w	X(w)
1/6	1 2 3	3
1/6	1 3 2	1
1/6	2 1 3	1
1/6	2 3 1	0
1/6	3 1 2	0
1/6	3 2 1	1

Expectation

The **expectation/expected value/average** of a discrete random variable X is

$$E[X] = \sum_{\omega \in \Omega} X(\omega) P(\omega)$$

Or equivalently,

$$E[X] = \sum_{k \in \Omega_X} k \cdot p_X(k)$$

The interpretation is that we take an average of the values in Ω_X, but weighted by their probabilities.

Homeworks of 3 students returned randomly

- Each permutation equally likely
- X: # people who get their own homework
- What is E(X)?

Prob	Outcome w	X(w)
1/6	1 2 3	3
1/6	1 3 2	1
1/6	2 1 3	1
1/6	2 3 1	0
1/6	3 1 2	0
1/6	3 2 1	1

Repeated coin flipping

Flip a biased coin with probability p of coming up Heads n times. Each flip independent of all others.

X is number of Heads.

What is E(X)?

Repeated coin flipping

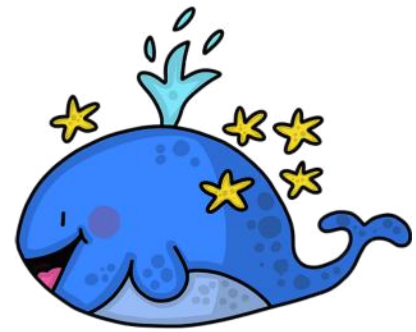
Flip a biased coin with probability p of coming up Heads n times.

X is number of Heads.

What is E(X)?

$$
\begin{aligned}
E[X] &= \sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} i \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} n \binom{n-1}{i-1} p^i (1-p)^{n-i} \\
&= np \sum_{i=1}^{n} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i} \\
&= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} \\
&= np(p + (1-p))^{n-1} = np
\end{aligned}
$$

Linearity of Expectation (Idea)



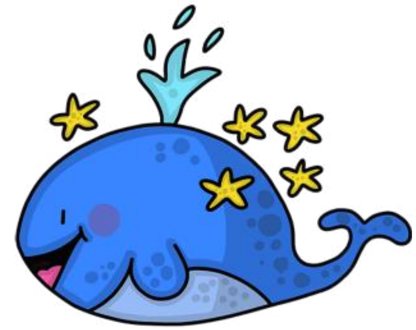
Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with $E[X] = 3$.
- Every day your friend catches Y fish, with $E[Y] = 7$.

How many fish do the two of you bring in ($Z = X + Y$) on an average day?

$$E[Z] = E[X + Y] =$$

Linearity of Expectation (Idea)



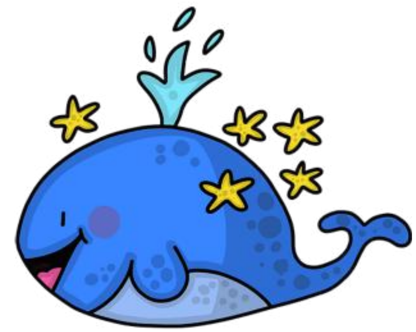
Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with **E[X]** = 3.
- Every day your friend catches **Y** fish, with **E[Y]** = 7.

How many fish do the two of you bring in ($Z = X + Y$) on an average day?

$E[Z] = E[X + Y] =$

Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

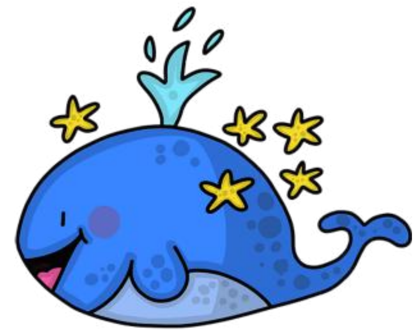
- Every day you catch X fish, with E[X] = 3.
- Every day your friend catches Y fish, with E[Y] = 7.

How many fish do the two of you bring in ($Z = X + Y$) on an average day?

$$E[Z] = E[X + Y] = E[X] + E[Y] = 3 + 7 = 10$$

You can sell each fish for $5 at a store, but you need to pay $20 in rent. How much profit do you expect to make? $E[5Z - 20] =$

Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with E[X] = 3.
- Every day your friend catches Y fish, with E[Y] = 7.

How many fish do the two of you bring in (Z = X + Y) on an average day?

$$E[Z] = E[X + Y] = E[X] + E[Y] = 3 + 7 = 10$$

You can sell each fish for \$5 at a store, but you need to pay \$20 in rent. How much profit do you expect to make? $E[5Z - 20] = 5E[Z] - 20 = 5 \times 10 - 20 = 30$

Linearity of Expectation (LoE)

Linearity of Expectation: Let Ω be the sample space of an experiment, $X, Y: \Omega \to \mathbb{R}$ be (possibly "dependent") random variables both defined on Ω, and $a, b, c \in \mathbb{R}$ be scalars. Then,

$$E[X + Y] = E[X] + E[Y]$$

and

$$E[aX + b] = aE[X] + b$$

Combining them gives,

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$

Linearity of Expectation (Proof)

$$E[X] + E[Y] = \sum_{\omega\in\Omega} X(\omega)P(\omega) + \sum_{\omega\in\Omega} Y(\omega)P(\omega)$$

$$= \sum_{\omega\in\Omega} \big(X(\omega) + Y(\omega)\big)P(\omega)$$

$$= \sum_{\omega\in\Omega} (X + Y)(\omega)\, P(\omega)$$

$$= E[X + Y]$$

Corollary: linearity for sum of lots of r.v.s

$$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$$

Proof by induction!

Homeworks of students returned randomly

- Each permutation equally likely
- X: # people who get their own homework
- What is E(X) when there are n students?

Prob	Outcome w	X			
1/6	1 2 3	3			
1/6	1 3 2	1			
1/6	2 1 3	1			
1/6	2 3 1	0			
1/6	3 1 2	0			
1/6	3 2 1	1			

Indicator random variable

- For any event A, can define the indicator random variable for A

Computing complicated expectations

- Often boils down to finding the right way to decompose the random variable into simple random variables (often indicator random variables) and then applying linearity of expectation.

Repeated coin flipping

Flip a biased coin with probability p of coming up Heads n times.

X is number of Heads.

What is E(X)?

$$\begin{aligned}
E[X] &= \sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} i \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} n \binom{n-1}{i-1} p^i (1-p)^{n-i} \\
&= np \sum_{i=1}^{n} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i} \\
&= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} \\
&= np(p + (1-p))^{n-1} = np
\end{aligned}$$

Repeated coin flipping

Flip a biased coin with probability p of coming up Heads n times.

X is number of Heads.

What is E(X)?

Pairs with same birthday

- In a class of m students, on average how many pairs of people have the same birthday?

Pairs with same birthday

- In a class of m students, on average how many pairs of people have the same birthday?

Rotating the table

n people are sitting around a circular table.
There is a nametag in each place
Nobody is sitting in front of their own nametag.
Rotate the table by a random number k of positions between 1 and n-1 (equally likely).
X is the number of people that end up front of their own nametag.
What is E(X)?

Linearity of Expectation with Indicators

For an indicator RV X_i,

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1)$$

Linearity is special!

- In general $E(g(X)) \neq g(E(X))$

E[X•Y] = E[X] • E[Y]
E[X²] = E[X]²
E[X/Y] = E[X] / E[Y]
E[asinh(X)] = asinh(E[X])
•
•
•

$$X = \begin{cases} 1 & \text{with prob } 1/2 \\ -1 & \text{with prob } 1/2 \end{cases}$$

Linearity is special!

- In general $E(g(X)) \neq g(E(X))$
- How DO we compute $E(g(X))$?

Homeworks of 3 students returned randomly

- Each permutation equally likely
- X: # people who get their own homework
- What is $E(X^3 \bmod 2)$?

Prob	Outcome w	X(w)
1/6	1 2 3	3
1/6	1 3 2	1
1/6	2 1 3	1
1/6	2 3 1	0
1/6	3 1 2	0
1/6	3 2 1	1

Law of the Unconscious Statistician (LOTUS)

Law of the Unconscious Statistician (LOTUS): Let X be a discrete random variable with range Ω_X and $g: D \to \mathbb{R}$ be a function defined at least over Ω_X ($\Omega_X \subseteq D$). Then,

$$E[g(X)] = \sum_{b \in \Omega_X} g(b) p_X(b)$$

Note that in general, $E[g(X)] \neq g(E[X])$. For example, $E[X^2] \neq (E[X])^2$, or $E[\log(X)] \neq \log(E[X])$.

Probability

3.3 Variance and standard Deviation

Most slides by Alex Tsun

Variance (Intuition)

Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me $1.
- If Tails, I pay you $1.

Game 2:

- If Heads, you pay me $1000.
- If Tails, I pay you $1000.

Variance (Intuition)

HOW FAR IS A RANDOM VARIABLE FROM ITS MEAN, ON AVERAGE?

$$X - E[X]$$

Variance (Intuition)

How far is a random variable from its mean, on average?

$$| X - E[X] |$$

Variance (Intuition)

How far is a random variable from its mean, on average?

$$E[\,|\,X - E[X]\,|\,]$$

$$E[\,(\,X - E[X]\,)^2\,]$$

Variance and Standard Deviation (SD)

Variance: The variance of a random variable X is

$$Var(X) = E[(X - E[X])^2]$$

The variance is always nonnegative since we take an expectation of a nonnegative random variable $(X - E[X])^2$. We can also show that for any scalars $a, b \in \mathbb{R}$,

$$Var(aX + b) = a^2 \, Var(X)$$

Variance and Standard Deviation (SD)

Variance: The variance of a random variable X is

More Useful

$$Var(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

The variance is always nonnegative since we take an expectation of a nonnegative random variable $(X - E[X])^2$. We can also show that for any scalars $a, b \in \mathbb{R}$,

$$Var(aX + b) = a^2\, Var(X)$$

Variance and Standard Deviation (SD)

Variance: The variance of a random variable X is

More Useful

$$Var(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

The variance is always nonnegative since we take an expectation of a nonnegative random variable $(X - E[X])^2$. We can also show that for any scalars $a, b \in \mathbb{R}$,

$$Var(aX + b) = a^2 \, Var(X)$$

Standard Deviation (SD): The standard deviation of a random variable X is

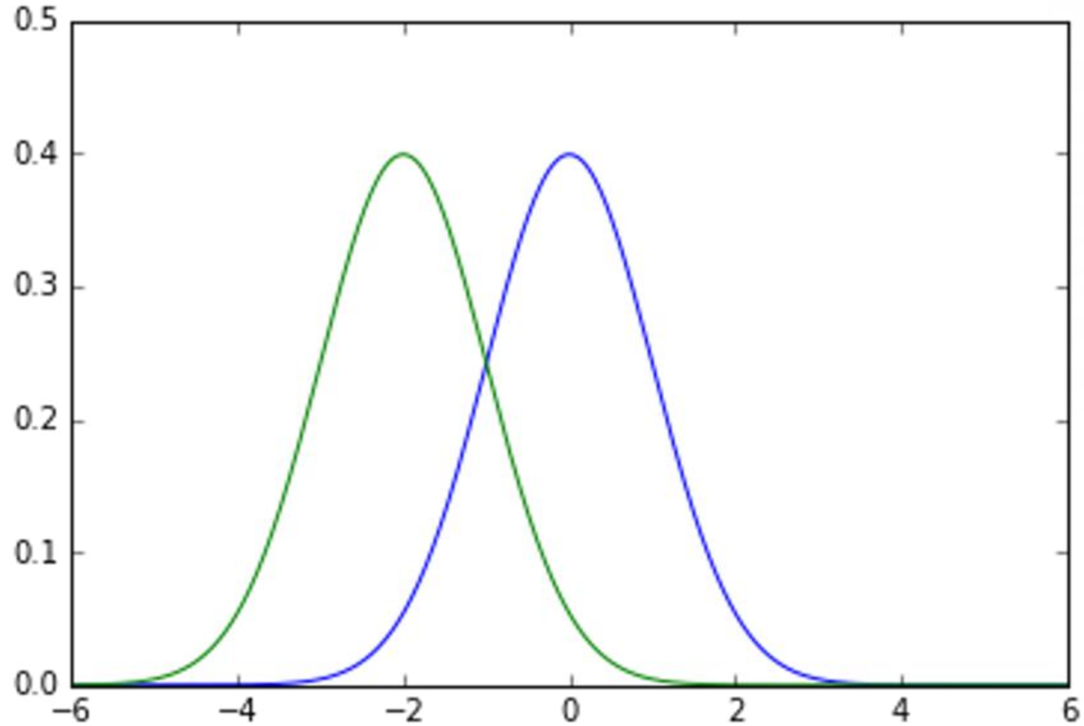
$$\sigma_X = \sqrt{Var(X)}$$

We want this because the units of variance are squared in terms of the original variable X, and this "undo's" our squaring, returning the units to the same as X.

Variance (Property)

$$Var(aX + b) = a^2 Var(X)$$

$$Var(X + b) = Var(X)$$



Variance (Property)

$$Var(aX + b) = a^2Var(X)$$

$$Var(aX) = E[(aX)^2] - (E[aX])^2 = E[a^2X^2] - (aE[X])^2$$

$$= a^2E[X^2] - a^2(E[X])^2 = a^2(E[X^2] - E[X]^2) = a^2Var(X)$$

Variance

Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me $1.
- If Tails, I pay you $1.

Game 2:

- If Heads, you pay me $1000.
- If Tails, I pay you $1000.

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$$

$$E[X^2] = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + \cdots + 6^2 \cdot \frac{1}{6} = \frac{91}{6}$$

LOTUS

$$Var(X) = E[X^2] - E[X]^2 = \frac{91}{6} - (3.5)^2 = \frac{35}{12}$$

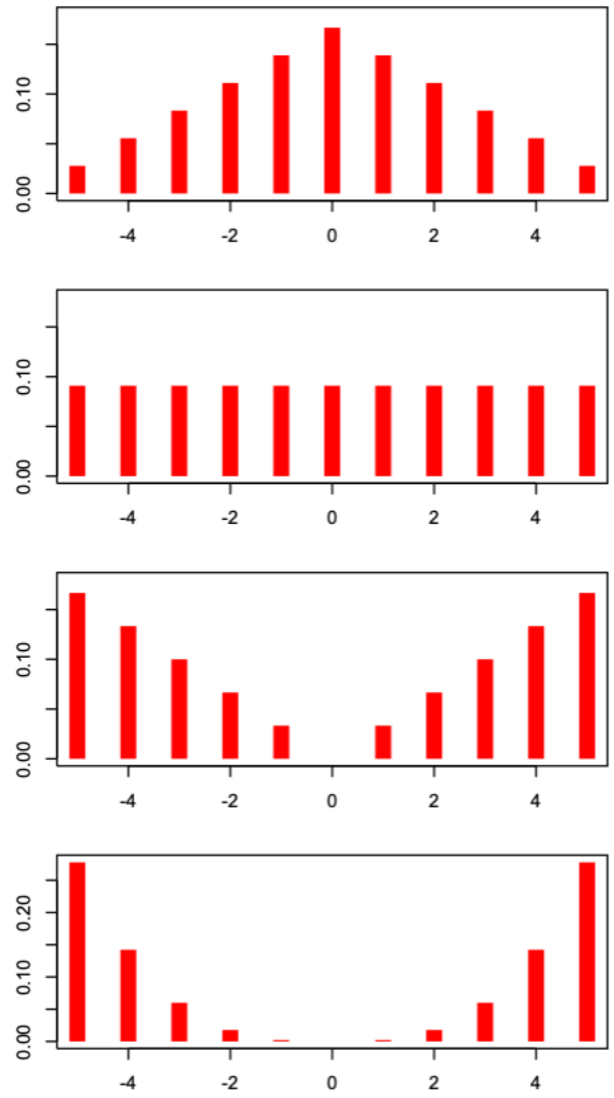
In general, var(X+Y) $\neq$ Var(X) + Var(Y)

Example 1:

- $X = \pm 1$ each with prob ½
- E(X)= ? Var(X) = ?
- How about Y = -X

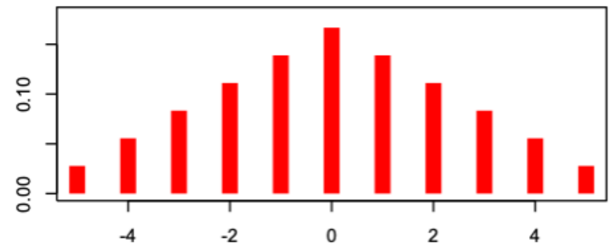
Example 2: What is Var(X+X)?

Variance in pictures

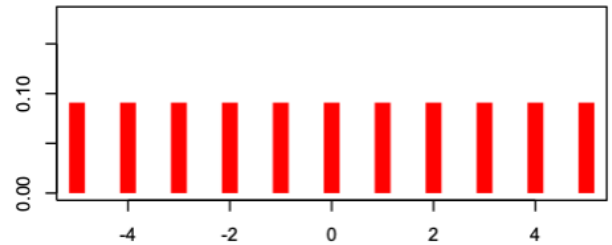


Variance in pictures

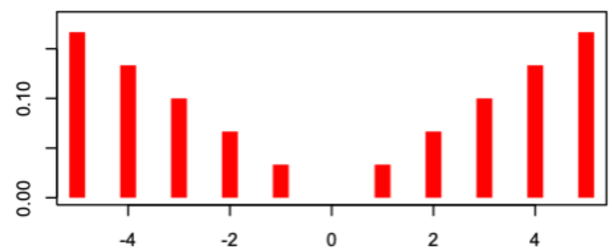
$\sigma^2 = 5.83$



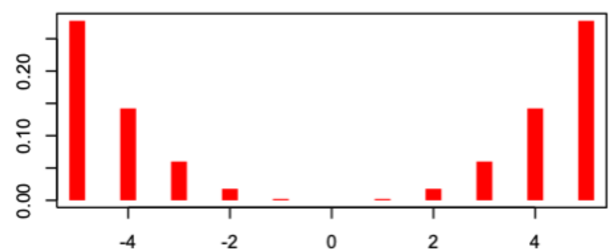
$\sigma^2 = 10$



$\sigma^2 = 15$



$\sigma^2 = 19.7$



One more linearity of expectation practice problem

Given a DNA sequence of length n

e.g. AAATGAATGAATCC……

where each position is

A with probability p_A

T with probability p_T

G with probability p_G

C with probability p_C.

What is the expected number of occurrences of the substring AATGAAT?

AAATGAATGAATCC AAATGAATGAATCC

Random Picture