CSC380: Principles of Data Science

Probability Primer 6

Xinchen Yu

Acknowledgement: Built on Jason Pacheco, Kwang-Sung Jun, Chicheng Zhang's slides

Announcements

Participation

- A total of 10 points (10% of final grade).
- Ask questions in-person in / after the class: +1 point
    - Don’t forget to let me know your names after class :)
- Answer 1 question in the lecture/on piazza: +1 point

HW1 solution out in D2L -> Content

HW2 due Sep 15, this Friday 11:59pm

Review

- Expectation

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

- Properties

$$E[X + Y] = E[X] + E[Y]$$
$$E[cX] = cE[X]$$
$$E[c] = c$$

c is a constant

- Conditional expected value

$$E[X|Y = y] = \sum_x x \cdot p(X = x|Y = y)$$

- Variance

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

- Properties

$$Var[cX] = c^2 Var[X]$$

- Covariance

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$
$$= E[XY] - E[X]E[Y]$$

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

- Variance of $X + Y$

$$Var[X + Y] = Var[X] + Var[Y] + 2Cov[X, Y]$$

Outline

- For independent RVs X_1 and X_2
  - $E(X_1 X_2)$
  - $Var(X_1 + X_2)$
  - $Cov(X_{1,} X_2)$

Independence and Moments

Theorem: *If* $X \perp Y$ *then* $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$.

Comparison: $\mathbf{E}[X+Y] = \mathbf{E}[X] + \mathbf{E}[Y]$ regardless of independence!

Independence and Moments

Scaling of Summations

$$\lambda \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \lambda x_i$$

Theorem: *If* $X \perp Y$ *then* $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$.

Proof:

$$\mathbf{E}[XY] = \sum_x \sum_y (x \cdot y) p(X = x, Y = y)$$

$$= \sum_x \sum_y (x \cdot y) p(X = x) p(Y = y) \qquad \textbf{(Independence)}$$

$$= \left(\sum_x x \cdot p(X = x) \right) \left(\sum_y y \cdot p(Y = y) \right) = \mathbf{E}[X]\mathbf{E}[Y] \qquad \textbf{(Linearity of Sum)}$$

Example *Let* $X_1, X_2 \in \{1, \ldots, 6\}$ *be RVs representing the result of rolling two fair standard dice.* *What is the mean of their product?*

$$\mathbf{E}[X_1 X_2] = \mathbf{E}[X_1]\mathbf{E}[X_2] = 3.5^2 = 12.25$$

Independence and Moments

Question: *What is the variance of their sum (recall independence)?*

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$
$$= E[XY] - E[X]E[Y]$$

- Proof 1:

$$\begin{aligned}
\mathbf{Var}[X_1 + X_2] &= \mathbf{Var}[X_1] + \mathbf{Var}[X_2] + 2\mathbf{Cov}(X_1, X_2) \\
&= \mathbf{Var}[X_1] + \mathbf{Var}[X_2] + 2\mathbf{E}[(X_1 - \mathbf{E}[X_1])(X_2 - \mathbf{E}[X_2])] \\
&= \mathbf{Var}[X_1] + \mathbf{Var}[X_2] + 2\mathbf{E}[(X_1 - \mathbf{E}[X_1])]\mathbf{E}[(X_2 - \mathbf{E}[X_2])] \\
&= \mathbf{Var}[X_1] + \mathbf{Var}[X_2] + 2\left(\mathbf{E}[X_1] - \mathbf{E}[X_1]\right)\left(\mathbf{E}[X_2] - \mathbf{E}[X_2]\right) \\
&= \mathbf{Var}[X_1] + \mathbf{Var}[X_2]
\end{aligned}$$

- $A \perp B \Rightarrow f(A) \perp f(B)$
- $f(\mathrm{X}) = \mathrm{X} - \mathrm{E}[\mathrm{X}]$
- $\mathrm{E}[\mathrm{f(A)f(B)}] = \mathrm{E}[\mathrm{f(A)}]\mathrm{E}[\mathrm{f(B)}]$

- Proof 2:

$$\begin{aligned}
Var[X_1 + X_2] &= Var[X_1] + Var[X_2] + 2Cov[X_1, X_2] \\
&= Var[X_1] + Var[X_2] + 2(E[X_1X_2] - E[X_1]E[X_2]) \\
&= Var[X_1] + Var[X_2] + 2(E[X_1]E[X_2] - E[X_1]E[X_2]) \\
&= Var[X_1] + Var[X_2]
\end{aligned}$$

Independence and Moments

Recall that for any two RVs X and Y variance is not a linear function,

$$\mathbf{Var}[X+Y] = \mathbf{Var}[X] + \mathbf{Var}[Y] + 2\mathbf{Cov}(X,Y)$$

<u>If X and Y are independent then they have zero covariance</u>,

$$\mathbf{Cov}(X,Y) = 0$$

Thus, $$\mathbf{Var}[X+Y] = \mathbf{Var}[X] + \mathbf{Var}[Y]$$

And, for a collection of independent RVs $X_1, X_2, \ldots, X_N$ we have,

$$\mathbf{Var}(\sum_{i=1}^{N} X_i) = \sum_{i=1}^{N} \mathbf{Var}(X_i)$$

Q: Is variance a linear operator under independence?

A: No! $Var(cX) \neq c\,Var(X)$ for a constant c. Rather, $Var(cX) = c^2 Var(X)$.

Linearity

In mathematics, a linear map or linear function $f(x)$ is a function that satisfies the two properties:[1]

- Additivity: $f(x + y) = f(x) + f(y)$.
- Homogeneity of degree 1: $f(\alpha x) = \alpha f(x)$ for all α. Homogeneous must pass: $f(zx, zy) = z^n f(x, y)$

Homogeneous?

$f(x, y) = 4x^2 + y^2 \Rightarrow$ homogeneous with degree 2: $f(zx, zy) = z^2 f(x, y)$

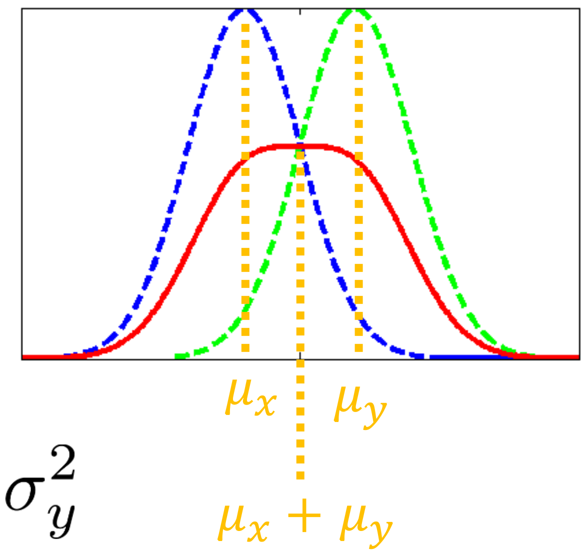
$\Rightarrow$ not linear

So, expectation is a linear function/operator, but variance is not !

We will just say ”linearity of expectation”

Example: Independent Gaussian RVs

Let X and Y be **independent** Gaussian RV with,



$$X \sim \mathcal{N}(\mu_x, \sigma_x^2) \qquad Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$$

(Property of Gaussian: $\mathrm{E}[X] = \mu_x$, $Var[X] = \sigma_x^2$)

What is the variance of their sum?

$$\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y) = \sigma_x^2 + \sigma_y^2$$

What is the mean of their product?

$$\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y] = \mu_x \mu_y$$

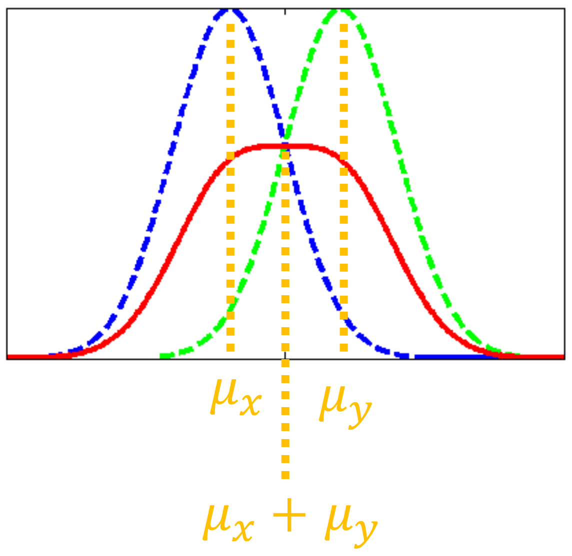
Suppose X and Y are **dependent**, what is the mean of their sum?

$$\mathbf{E}[X+Y] = \mathbf{E}[X] + \mathbf{E}[Y] = \mu_x + \mu_y$$

The amazing Gaussian

Let X and Y be **independent** Gaussian RVs with,

$$X \sim \mathcal{N}(\mu_x, \sigma_x^2) \qquad\qquad Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$$



For normal distributions

- Closed under additivity:

$$X \sim \mathcal{N}(\mu_x, \sigma_x^2) \quad Y \sim \mathcal{N}(\mu_y, \sigma_y^2) \quad , X \perp Y$$

$$X + Y \sim \mathcal{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$$

- Closed under affine transformation (a and b constant):

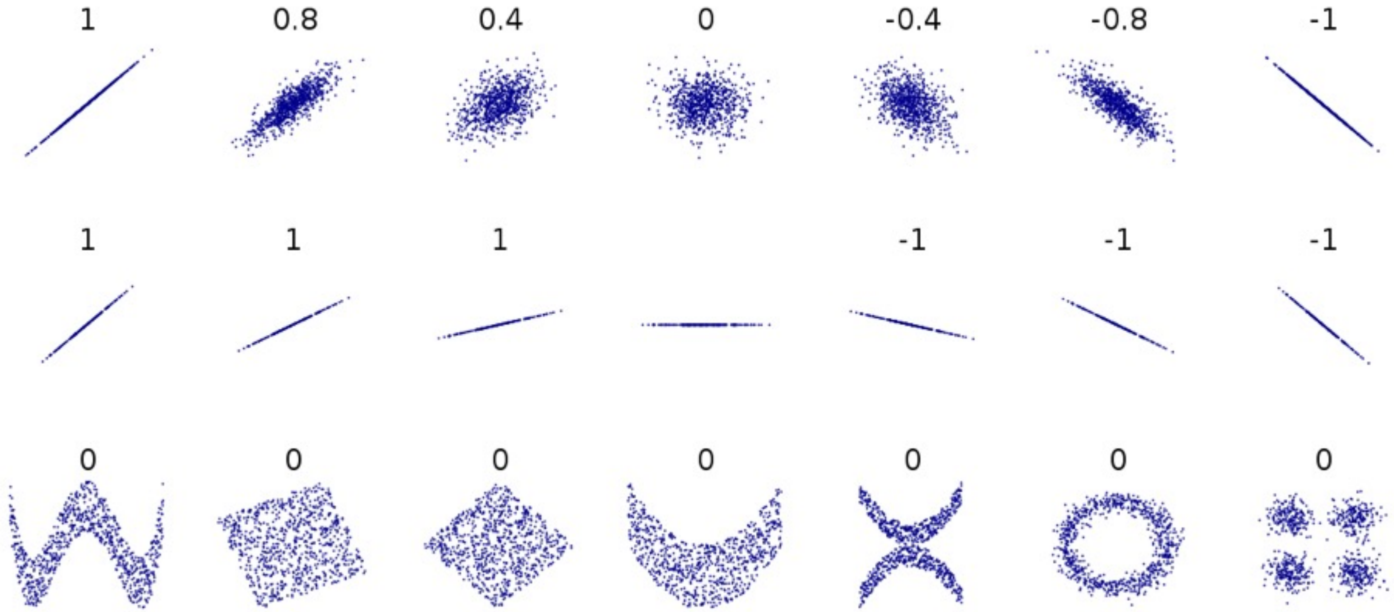
$$aX + b \sim \mathcal{N}(a\mu_x + b, a^2\sigma_x^2)$$

Independence and Moments

On slide page 6, If X and Y are independent RVs, then:

$$\mathbf{Cov}(X, Y) = 0$$

The reverse is not true! $(\mathbf{Cov}(X, Y) = 0) \nRightarrow X \perp Y$



Counter Example

- Let X, Z be independent RV that is −1 or +1 with probability 1/2.
- Let $Y = Z \cdot I\{X = 1\}$
- Claim: $Cov(X, Y) = 0$ but X and Y are dependent.

Indicator function:
$I\{X = 1\} = 1, if\ X = 1$
$I\{X = 1\} = 0, if\ X \neq 1$

$Cov(X, Y) = E[XY] - E[X]E[Y]$

	X	Z	Y	XY
1	½	½	¼	¼
-1	½	½	¼	¼
0	N/A	N/A	½	½

$Y = Z \cdot 1, if\ X = 1, \qquad Y = Z \cdot 0, if\ X = -1$

$$P(Y = 1) = P(X = 1, Z = 1) = P(X = 1) \cdot P(Z = 1) = \frac{1}{4}$$

$$P(Y = -1) = P(X = 1, Z = -1) = \frac{1}{4}$$

$$P(Y = 0) = P(X = -1) = \frac{1}{2}$$

$$P(XY = 1) = P(X = 1, Y = 1) + P(X = -1, Y = -1)$$
$$= P(X = 1, Z = 1) + 0 = \frac{1}{4}$$

$$P(XY = -1) = P(X = 1, Y = -1) + P(X = -1, Y = 1)$$
$$= P(X = 1, Z = -1) + 0 = \frac{1}{4}$$

$$P(XY = 0) = P(Y = 0) = \frac{1}{2}$$

$\mathrm{E[X]}=1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0$
$\mathrm{E[Y]}=0$
$\mathrm{E[XY]}=0$

$P(X = 1, Y = 0) = 0$

$\mathrm{P(X = 1)P(Y = 0)} = \frac{1}{4} \qquad 0 \neq \frac{1}{4}!$

Moments of Continuous RVs

Replace all sums with integrals,

$$\mathbf{E}[X] = \int x p(x)\, dx \qquad \mathbf{Var}[X] = \int (x - \mathbf{E}[X])^2 p(x)\, dx$$

- All properties push through, as you would expect (e.g. law of total expectation, conditional expectation, etc.)

(and use PDF p(x) instead of PMF P(X=x))

Exercise

<u>Question:</u> *Roll two dice and let their outcomes be* $X_1, X_2 \in \{1, \ldots, 6\}$ *for die 1 and die 2, respectively. Recall the definition of conditional probability,*

$$p(X_1 \mid X_2) = \frac{p(X_1, X_2)}{p(X_2)}$$

Which of the following are true?

a) $p(X_1 = 1 | X_2 = 1) > p(X_1 = 1)$

b) $\boxed{p(X_1 = 1 | X_2 = 1) = p(X_1 = 1)}$ Outcome of die 2 doesn't *affect* die 1

c) $p(X_1 = 1 | X_2 = 1) < p(X_1 = 1)$

Exercise

Question: *Let* $X_1 \in \{1, \ldots, 6\}$ *be outcome of die 1, as before. Now let* $X_3 \in \{2, 3, \ldots, 12\}$ *be the* sum of both dice. Which of the following are true?

a) $p(X_1 = 1 | X_3 = 3) > p(X_1 = 1)$

b) $p(X_1 = 1 | X_3 = 3) = p(X_1 = 1)$

c) $p(X_1 = 1 | X_3 = 3) < p(X_1 = 1)$

Only 2 ways to get $X_3 = 3$ *, each with equal probability:*

$$(X_1 = 1, X_2 = 2) \quad \text{or} \quad (X_1 = 2, X_2 = 1)$$

so

$$p(X_1 = 1 \mid X_3 = 3) = \frac{1}{2} > \frac{1}{6} = p(X_1 = 1)$$

Review

We have covered a lot of ground on probability in short time…

Discrete Random Processes

- Definition of sample space / random events
- Axioms of probability
- Uniform probability of random event
- Random Variables
- Fundamental rules of probability (chain rule, conditional, law of total probability)

Probability Distributions

- Useful discrete probability mass functions
- Introduction to continuous probability
- Useful probability density functions

Moments / Independence

- Expected Value
- Linearity
- Variance, Covariance, Corr.
- Dependent / Independent RVs

Homework 2

- d) Assume I take the test twice, and receive a positive result in the first test and a negative result in the second test. Assume that the two test results are conditionally independent given the existence of the antibody. What is the probability that I have COVID-19 antibodies according to Bayes' rule?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

$S: antibody\ state$
$R_1: result\ of\ first\ test$
$R_2: result\ of\ second\ test$

$$P(S = T|R_1 = T, R_2 = F) = \frac{P(R_1 = T, R_2 = F|S = T)P(S = T)}{P(R_1 = T, R_2 = F)}$$

Law of total probability

$$\begin{aligned}
&P(R_1 = T, R_2 = F) \\
&= P(R_1 = T, R_2 = F, S = T) + P(R_1 = T, R_2 = F, S = F) \\
&= P(R_1 = T, R_2 = F|S = T)P(S = T) + P(R_1 = T, R_2 = F|S = F)P(S = F) \\
&= P(R_1 = T|S = T)P(R_2 = F|S = T)P(S = T) + P(R_1 = T|S = F)P(R_2 = F|S = F)P(S = F)
\end{aligned}$$