# DCP 1206: Probability
# Lecture 15 — Expected Value of Two Random Variables and Covariance

Ping-Chun Hsieh

November 8, 2019

# Announcements

- Midterm course evaluation (期中意見調查)
  - Google form: shorturl.at/ixAFP (posted on E3)
  - 11/8 (Friday) ～ 11/13 (Wednesday)

# Unboxing!

- What will we learn in the 2nd half of the course?

# 1. Sum of Independent Random Variables

Example: Toss "moon blocks"


- 3 possible outcomes: Yes / No / Laughing
- $p = P(\text{outcome is "Yes"})$
- Each toss is independent from other tosses

Question: How to learn $p$?

- Idea: Try $N$ tosses and take empirical average
  - Law of Large Numbers: empirical average = $p$, when $N \to \infty$
  - Concentration inequalities: empirical average $\approx p$, when $N$ large

# 2. Statistical Inference

Example: Toss “moon blocks” 


- 3 possible outcomes: Yes / No / Laughing
- $p = P(\text{outcome is “Yes”})$
- Each toss is <u>independent</u> from other tosses

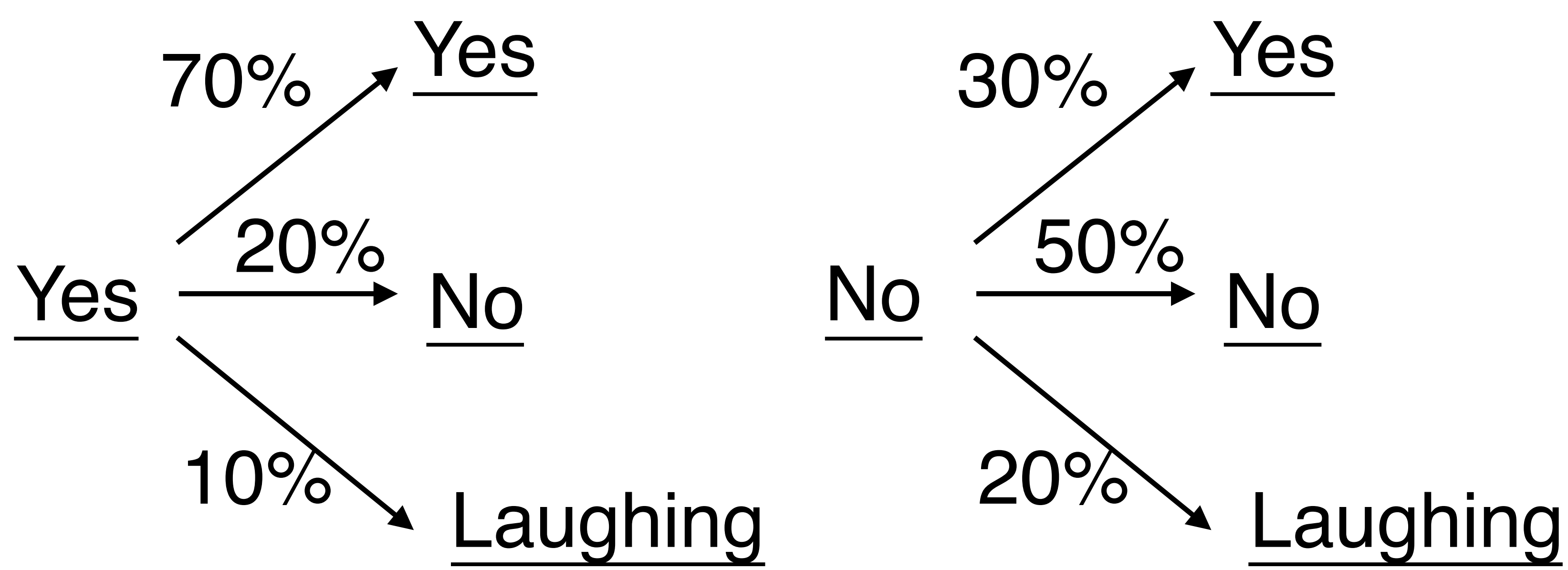
Question: Estimate $p$ with only 10 tosses?

- <u>2 Major Approaches</u>:
  - Maximum Likelihood Estimation
  - Bayesian Inference (e.g. HW1 Problem 6)
    - Learn when Beta & Gamma distributions are useful

# 3. Markov Chain and Beyond

Example: Toss “moon blocks” 


- 3 possible outcomes: Yes / No / Laughing
- Each toss depends on the previous toss (called “**Markov**”)




Question: What happened after many tosses?

- Idea: We will study “steady-state” behavior

# Quick Review

- Given 2 random variables $X, Y$: What have we learned so far?
  1. Joint CDF
  2. Marginal CDF
  3. Joint PMF / PDF
  4. Marginal PMF / PDF
  5. Independence

- Anything else?

# Quick Review

- Given 2 random variables $X, Y$: what have we learned so far?
  1. Joint CDF
  2. Marginal CDF
  3. Joint PMF / PDF
  4. Marginal PMF / PDF
  5. Independence
  6. Expected value involving both $X, Y$
  7. Covariance and correlation
  8. Conditional distribution
  9. Distribution of $X + Y$

# This Lecture

1. Expected Value Regarding 2 Random Variables

2. Covariance

- Reading material: Chapter 8.2 and 10.2

# 1. Expected Value Regarding Two Random Variables

# Recall: LOTUS for 1 Discrete Random Variable

**Expected Value of a Function of Discrete R.V.**:

1. Let $X$ be a discrete random variable with

- the set of possible values $S$
- PMF of $X$ is $p_X(x)$

2. Let $g(\cdot)$ be a real-valued function

The expectation of $g(X)$ is

$$E[g(X)] = \sum_{x \in S} g(x) \cdot p_X(x)$$

# LOTUS for 2 Discrete Random Variables

**Expected Value of a Function of 2 Discrete RVs**:

1. Let $X, Y$ be 2 discrete random variables with sets of possible values $S_X, S_Y$ and joint PMF $p(x, y)$

2. Let $g(\cdot\,,\cdot)$ be a function from $\mathbb{R}^2 \to \mathbb{R}$

The expected value of $g(X, Y)$ is

PMF

$$E[g(X,Y)] = \sum_{\substack{x \in S_X \\ y \in S_Y}} \big( g(x,y) \big)\big( p(x,y) \big)$$

$$= \sum_{x \in S_X} \sum_{y \in S_Y} g(x,y) \cdot p(x,y)$$

# Example: Using Joint PMF to Find Expected Value

- Example: Bus #2 (NCTU - Mackay - Train Station)
  - $X =$ traveling time from NCTU to Mackay
  - $Y =$ traveling time from Mackay to Train Station
  - $E[X + Y] = ?$

| Joint PMF | X=10 | X=15 | X=20 |
|---|---|---|---|
| Y=10 | 0.1 | 0.1 | 0.05 |
| Y=15 | 0.1 | 0.3 | 0.1 |
| Y=20 | 0.05 | 0.1 | 0.1 |

$$E[\underbrace{X+Y}_{g(X,Y)}] = \sum_{\text{all}(x,y)} (x+y)(p(x,y))$$

$$= (20)(0.1) + (25)(0.1) + (30)(0.05)$$

$$+ (25)(0.1) + (30)(0.3) + (35)(0.1)$$

$$+ (30)(0.05) + (35)(0.1) + (40)(0.1)$$

# Recall: LOTUS for 1 Continuous Random Variable

**Expected Value of a Function of 1 <u>Continuous</u> RV**:

Let $X$ be a <u>continuous</u> random variable with a PDF $f_X(x)$. Let $g(\,\cdot\,)$ be a real-valued function. Then,

$$E[g(X)] := \int_{-\infty}^{+\infty} \underbrace{g(x)}_{\text{function value}} \cdot \underbrace{f_X(x)}_{\text{PDF}} dx$$

# LOTUS for 2 Continuous Random Variables

**Expected Value of a Function of 2 Continuous RVs**:

1. Let $X, Y$ be 2 continuous random variables with joint PDF $f_{XY}(x, y)$
2. Let $g(\cdot\,,\cdot)$ be a function from $\mathbb{R}^2 \to \mathbb{R}$

The expected value of $g(X, Y)$ is

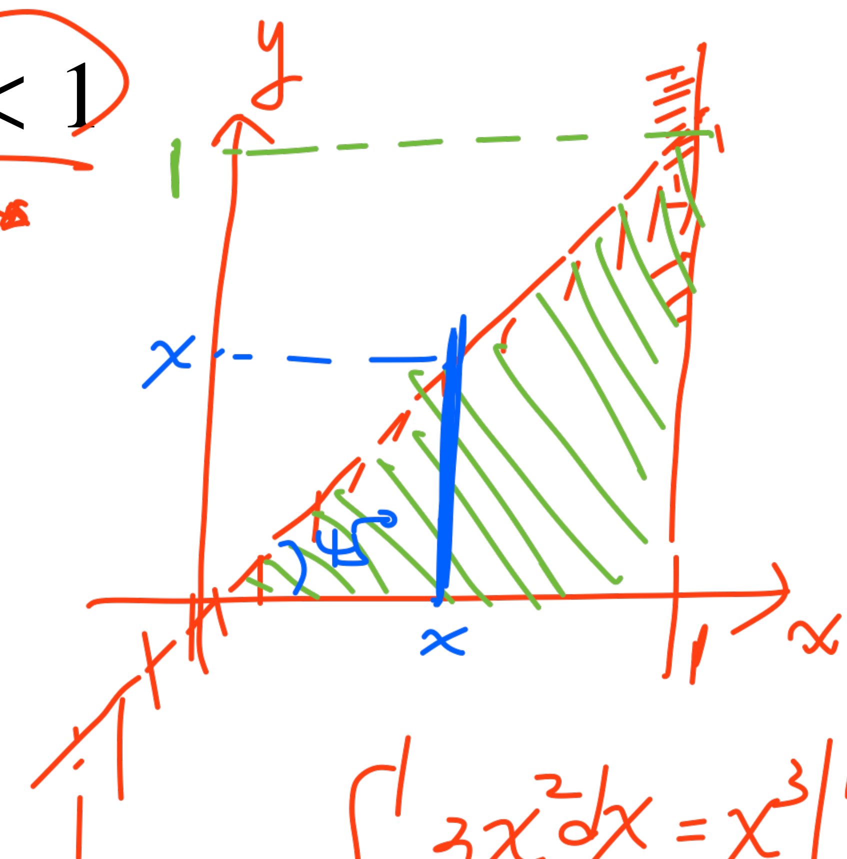
$$E[g(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \underbrace{(g(x,y))}_{\text{function value}} \underbrace{(f_{XY}(x,y))}_{\text{PDF}} \, dx\, dy$$

# Example: Using Joint PDF to Find Expected Value

- Example:

$$f(x,y) = \begin{cases} 2 & , \text{ if } 0 < y < x < 1 \\ 0 & , \text{ otherwise} \end{cases}$$

- $E[X+Y] = ?$



$$E[X+Y] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \underbrace{(x+y)}_{\text{function value}} \underbrace{(f(x,y))}_{\text{joint PDF}} \, dy\, dx$$

$$= \int\int (x+y)(2)\, dy\, dx$$

$$= \int_0^1 \int_0^x \underbrace{(x+y)\cdot 2}_{2x+2y}\, dy\, dx = \int_0^1 \left( 2xy + y^2 \Big|_0^x \right) dx$$

$$2x^2 + x^2 = 3x^2$$

$$\int_0^1 3x^2\, dx = x^3 \Big|_0^1 = 1$$

# Useful Property (I)

- **Linearity Property**:

$$E[\underline{\alpha \cdot g_1(X, Y)} + \underline{\beta \cdot g_2(X, Y)}] = \underline{\alpha E[g_1(X, Y)]} + \underline{\beta E[g_2(X, Y)]}$$

- Remark: $X, Y$ are <u>NOT</u> required to be independent
- Remark: This results holds for both discrete and continuous cases
- Proof:

$$E[\alpha \cdot g_1 + \beta \cdot g_2] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\underbrace{\alpha g_1 + \beta g_2}) (f(x,y))\, dx\, dy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \alpha g_1 \cdot f(x,y)\, dx\, dy = \alpha \cdot E[g_1(X,Y)]$$

$$+ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \beta g_2 \cdot f(x,y)\, dx\, dy = \beta \cdot E[g_2(X,Y)]$$

# A Corollary of Linearity Property

Theorem

Proposition ✓

Lemma

Definition

Axiom

Corollary

Remark

Conjecture

猜想.

**Corollary**: Let $X, Y$ be 2 random variables. Then,

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

- Remark: Binomial$(2,p)$ has an expected value of $2p$
- How about Binomial$(n,p)$?

Define $X \sim \text{Bernoulli}(p)$, $Y \sim \text{Bernoulli}(p)$; assuming $X, Y$ indep.

$Z = X + Y \sim \text{Binomial}(2,p)$

# Useful Property (II)

- **Property under independence**: Suppose $X, Y$ are independent random variables. Then, we have

$$E[g(X) \cdot h(Y)] = E[g(X)] \cdot E[h(Y)]$$

- Remark: This result holds for both discrete and continuous cases
- Proof:

# $E[XY] = E[X]E[Y]$ If $X, Y$ Are Independent

**Corollary**: Let $X, Y$ be 2 independent random variables. Then,

$$E[XY] = E[X]E[Y]$$

- Question: How about the reverse argument?

# $E[XY] = E[X]E[Y] \not\Rightarrow X, Y$ Independent

- Example: Let $X$ be a continuous uniform r.v. on $[-1,1]$.
  - Define $Y = -X$
  - $E[X] =$ ? $E[Y] =$ ?
  - $E[XY] =$ ?
  - Are $X, Y$ independent?

# Example: $X$, $Y$ Not Independent and $E[XY] \neq E[X]E[Y]$

- Example: Let $X$ be a <u>continuous uniform</u> r.v. on $[-1,1]$.
  - Define $Y = X$
  - Are $X$, $Y$ independent?
  - $E[X] =$ ? $E[Y] =$ ?
  - $E[XY] =$ ?

# When Do We Need $E[XY]$?

- Example: $\mathrm{Var}[X + Y]$

# More on $E[XY]$: Cauchy-Schwarz Inequality

- Recall: Cauchy Inequality in high school

# Cauchy-Schwarz Inequality

- **Cauchy-Schwarz Inequality**: Let $X$, $Y$ be two random variables. Then, we have

$$E[X^2] \cdot E[Y^2] \geq (E[XY])^2$$

- Question: Under what condition do we have "="?

# Proof of Cauchy-Schwarz Inequality

$$E[X^2] \cdot E[Y^2] \geq (E[XY])^2$$

- Hint: Start from that $E[(tX + Y)^2] \geq 0$
- Proof:

# 2. Covariance

# Motivating Example for “Covariance”

- Example: Bus #2 (NCTU - Mackay - Train Station)
  - $X =$ traveling time from NCTU to Mackay
  - $Y =$ traveling time from Mackay to Train Station
  - We want to know Var$[X + Y]$
  - Question: Given Var$[X]$ and Var$[Y]$, can we get Var$[X + Y]$?

# Covariance and Variance

- Property:

$$\text{Var}[aX + bY] = a^2\text{Var}[X] + b^2\text{Var}[Y] + 2abE[(X - E[X])(Y - E[Y])]$$

# Covariance (Formally)

- **Covariance**: Let $X, Y$ be two random variables. Then, the covariance of $X$ and $Y$ is defined as

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$$

- $\text{Cov}(X, Y) = 0$: $X, Y$ are said to be ________________
- $\text{Cov}(X, Y) > 0$: $X, Y$ are said to be ________________
- $\text{Cov}(X, Y) < 0$: $X, Y$ are said to be ________________
- Intuition:

# A Simplified Expression of Covariance

- Let $X, Y$ be two random variables. Then, the covariance of $X$ and $Y$ can also be written as

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

- Question: How to show this?

- Question: If $X, Y$ are independent, then $\text{Cov}(X, Y) = ?$
- Question: How about the reverse argument?

# Example: Uncorrelated $\nRightarrow$ Independence

- Example: The pair of random variables $(X, Y)$ takes the values $(1,0), (0,1), (-1,0), (0,-1)$, each with probability $\frac{1}{4}$
  - Cov$(X, Y)$ = ?
  - Are $X, Y$ independent?

# Example

- Example: Let $\theta$ be a continuous uniform random variable on $[0,2\pi]$. Define $X = \cos\theta$ and $Y = \sin\theta$
  - Cov$(X, Y)$ = ?
  - Are $X, Y$ independent?

# 1-Minute Summary

## 1. Expected Value Regarding 2 Random Variables

- Extend LOTUS for 2 random variables
- Independence: $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$
- Cauchy-Schwarz Inequality

## 2. Covariance

- Motivation / definition / simplified expression
- Uncorrelated vs Independence