STAT 3375Q: Introduction to Mathematical Statistics I

Lecture 17: Expected Value of a Function of Random Variables; Covariance; Conditional Expectation

Mary Lai Salvaña, Ph.D.

Department of Statistics
University of Connecticut

March 27, 2024

Outline

Previously...

Marginal Probability Distribution

- Let Y_1 and Y_2 be random variables with joint CDF $F(y_1, y_2)$
  - Marginal CDF of Y_1: $F_1(y_1) = F(y_1, \infty)$
  - Marginal CDF of Y_2: $F_2(y_2) = F(\infty, y_2)$
- Let Y_1 and Y_2 be jointly discrete random variables with joint PMF $p(y_1, y_2)$.
  - Marginal PMF of Y_1: $p_1(y_1) = \sum_{\text{all } y_2} p(y_1, y_2)$
  - Marginal PMF of Y_2: $p_2(y_2) = \sum_{\text{all } y_1} p(y_1, y_2)$
- Let Y_1 and Y_2 be jointly continuous random variables with joint PDF $f(y_1, y_2)$.
  - Marginal PDF of Y_1: $f_1(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_2$
  - Marginal PDF of Y_2: $f_2(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_1$

Conditional Probability Distribution

- Let Y_1 and Y_2 be jointly discrete random variables with joint PMF $p(y_1, y_2)$ and marginal PMFs $p_1(y_1)$ and $p_2(y_2)$.
  - Conditional PMF of Y_1 given $Y_2 = y_2$:
    $p(y_1|y_2) = P(Y_1 = y_1|Y_2 = y_2) = \frac{P(Y_1=y_1, Y_2=y_2)}{P(Y_2=y_2)} = \frac{p(y_1,y_2)}{p_2(y_2)}$,

  provided that $p_2(y_2) > 0$.

- Let Y_1 and Y_2 be jointly continuous random variables with joint PDF $f(y_1, y_2)$ and marginal PDFs $f_1(y_1)$ and $f_2(y_2)$.
  - Conditional PDF of Y_1 given $Y_2 = y_2$: $f(y_1|y_2) = \frac{f(y_1,y_2)}{f_2(y_2)}$,

  provided that $f_2(y_2) > 0$.

- Let Y_1 and Y_2 be jointly continuous random variables with joint PDF $f(y_1, y_2)$.
  - Conditional CDF of Y_1 given $Y_2 = y_2$: $F(y_1|y_2) = P(Y_1 \leq y_1|Y_2 = y_2)$.

Independent Random Variables

- Let Y_1 have CDF $F_1(y_1)$, Y_2 have CDF $F_2(y_2)$, and Y_1 and Y_2 have joint CDF $F(y_1, y_2)$.
  - Y_1 and Y_2 independent: $F(y_1, y_2) = F_1(y_1)F_2(y_2)$
- If Y_1 and Y_2 are not independent, then they are dependent.
- Let Y_1 and Y_2 be jointly discrete random variables with joint PMF $p(y_1, y_2)$ and marginal PMFs $p_1(y_1)$ and $p_2(y_2)$.
  - Y_1 and Y_2 independent: $p(y_1, y_2) = p_1(y_1)p_2(y_2)$
- Let Y_1 and Y_2 be jointly continuous random variables with joint PDF $f(y_1, y_2)$ and marginal PDFs $f_1(y_1)$ and $f_2(y_2)$.
  - Y_1 and Y_2 independent: $f(y_1, y_2) = f_1(y_1)f_2(y_2)$
- Useful Theorem for Independence: $f(y_1, y_2) = g(y_1)h(y_2)$, where $g(y_1)$ is a nonnegative function of y_1 alone and $h(y_2)$ is a nonnegative function of y_2 alone.

Expected Value of a Function of Random Variables

Expected Value of a Function of Random Variables

Definition: Expected Value of a Function of Discrete RVs

Let $g(Y_1, Y_2, \ldots, Y_n)$ be a function of the discrete random variables $Y_1, Y_2, \ldots, Y_n$ with PMF $p(y_1, y_2, \ldots, y_n)$. The *expected value* of $g(Y_1, Y_2, \ldots, Y_n)$ is

$$E\{g(Y_1, Y_2, \ldots, Y_n)\} = \sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} g(Y_1, Y_2, \ldots, Y_n) p(y_1, y_2, \ldots, y_n).$$

- We will need the formula above to compute one of the most popular measures of DEPENDENCE of RVs:

$$\begin{aligned} \text{Cov}(Y_1, Y_2) &= E\{(Y_1 - \mu_1)(Y_2 - \mu_2)\} \\ &= \sum_{y_1} \sum_{y_2} (y_1 - \mu_1)(y_2 - \mu_2) p(y_1, y_2) \end{aligned}$$

Expected Value of a Function of Random Variables

Definition: Expected Value of a Function of Continuous RVs

Let $g(Y_1, Y_2, \ldots, Y_n)$ be a function of the continuous random variables $Y_1, Y_2, \ldots, Y_n$ with PDF $f(y_1, y_2, \ldots, y_n)$. The *expected value* of $g(Y_1, Y_2, \ldots, Y_n)$ is

$$E\{g(Y_1, Y_2, \ldots, Y_n)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(Y_1, Y_2, \ldots, Y_n) \times f(y_1, y_2, \ldots, y_n) dy_1 dy_2 \cdots dy_n.$$

- We will need the formula above to compute one of the most popular measures of DEPENDENCE of RVs:

$$\begin{aligned} \text{Cov}(Y_1, Y_2) &= E\{(Y_1 - \mu_1)(Y_2 - \mu_2)\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (y_1 - \mu_1)(y_2 - \mu_2) f(y_1, y_2) dy_1 dy_2 \end{aligned}$$

Expected Value of a Function of Random Variables

Theorem: Properties of Expected Value

1. Let c be a constant. Then, $E(c) = c$.

2. Let $g(Y_1, Y_2)$ be a function of Y_1 and Y_2 and let c be a constant. Then, $E\{cg(Y_1, Y_2)\} = cE\{g(Y_1, Y_2)\}$.

3. Let $g_1(Y_1, Y_2), g_2(Y_1, Y_2), \ldots, g_1k(Y_1, Y_2)$ be functions of Y_1 and Y_2. Then,

$$E\{g_1(Y_1, Y_2) + g_2(Y_1, Y_2) + \ldots + g_k(Y_1, Y_2)\}$$
$$= E\{g_1(Y_1, Y_2)\} + E\{g_2(Y_1, Y_2)\} + \quad \ldots \quad + E\{g_k(Y_1, Y_2)\}.$$

Expected Value of a Function of Random Variables

Theorem: Expected Value of Independent RVs

Let Y_1 and Y_2 be independent random variables (discrete or continuous) and $g(Y_1)$ and $h(Y_2)$ be functions of only Y_1 and Y_2, respectively. Then,

$$E\{g(Y_1)h(Y_2)\} = E\{g(Y_1)\}E\{h(Y_2)\},$$

provided that the expectations exist.

Covariance

Covariance

Definition: Covariance

If Y_1 and Y_2 are random variables with mean μ_1 and μ_2, respectively, the *covariance* of Y_1 and Y_2 is

$$\text{Cov}(Y_1, Y_2) = E\{(Y_1 - \mu_1)(Y_2 - \mu_2)\}.$$

- Covariance is a number quantifying average dependence between two random variables.
- Notation: $\sigma_{12} = \text{Cov}(Y_1, Y_2)$
- Discrete Case:
  $\text{Cov}(Y_1, Y_2) = \sum_{y_1} \sum_{y_2} (y_1 - \mu_1)(y_2 - \mu_2) p(y_1, y_2)$,

  Expected value of a function of discrete RV (Slide 8)

  where $p(y_1, y_2)$ is the joint PMF of Y_1 and Y_2.
- Continuous Case:
  $\text{Cov}(Y_1, Y_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y_1 - \mu_1)(y_2 - \mu_2) f(y_1, y_2) dy_1 dy_2$,

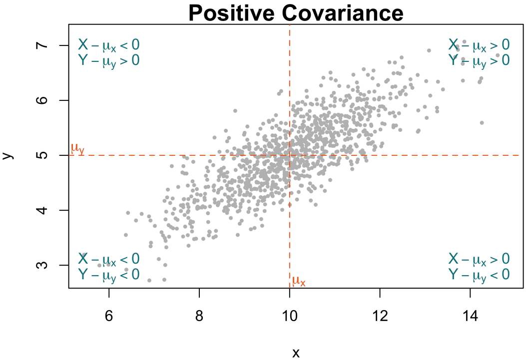
  Expected value of a function of continuous RV (Slide 9)

  where $f(y_1, y_2)$ is the joint PDF of Y_1 and Y_2.

Covariance: Sign of Covariance Reveals Relationship

$$\text{Cov}(X, Y) = E\{(X - \mu_x)(Y - \mu_y)\}$$

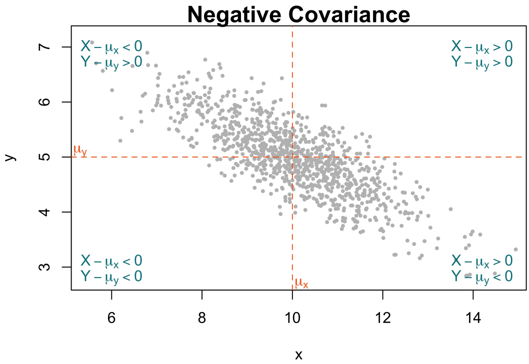
- $\text{Cov}(X, Y) > 0$ means a positive relationship between X and Y
  - When X increases, Y tends to increase.



Covariance: Sign of Covariance Reveals Relationship

$$\text{Cov}(X, Y) = E\{(X - \mu_x)(Y - \mu_y)\}$$

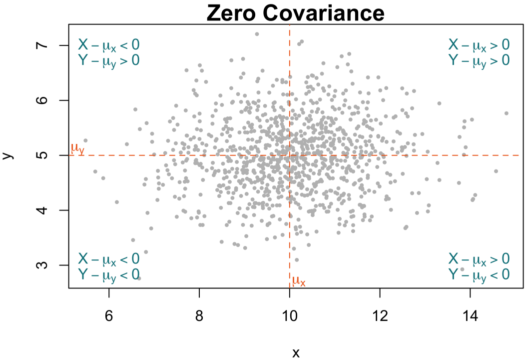
- $\text{Cov}(X, Y) < 0$ means a negative relationship between X and Y
  - When X increases, Y tends to decrease.



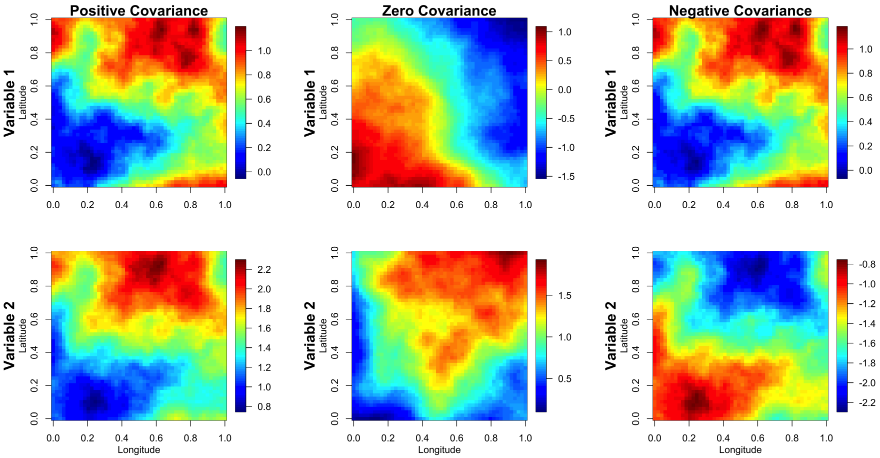
Covariance: Sign of Covariance Reveals Relationship

$$\text{Cov}(X, Y) = E\{(X - \mu_x)(Y - \mu_y)\}$$

- $\text{Cov}(X, Y) = 0$ means there is NO relationship between X and Y
  - When X increases, Y can increase or decrease.
  - We call X and Y uncorrelated random variables.



Covariance: Can Be Easily Detected



Covariance

Properties: More Useful Formula to Compute Covariance

1. If Y_1 and Y_2 are random variables with mean μ_1 and μ_2, respectively, then

$$\text{Cov}(Y_1, Y_2) = E(Y_1 Y_2) - \mu_1\mu_2.$$

Proof:

$$\begin{aligned}
\text{Cov}(Y_1, Y_2) &= E\{(Y_1 - \mu_1)(Y_2 - \mu_2)\} && \text{def'n of covariance} \\
&= E(Y_1 Y_2 - \mu_1 Y_2 - \mu_2 Y_1 + \mu_1\mu_2) \\
&= E(Y_1 Y_2) - E(\mu_1 Y_2) - E(\mu_2 Y_1) + E(\mu_1\mu_2) && \text{linearity of expectation} \\
&= E(Y_1 Y_2) - \mu_1 E(Y_2) - \mu_2 E(Y_1) + \mu_1\mu_2 && \text{expected value of a constant is itself} \\
&= E(Y_1 Y_2) - \mu_1\mu_2 - \mu_2\mu_1 + \mu_1\mu_2 && \text{def'n of expected value} \\
&= E(Y_1 Y_2) - \mu_1\mu_2. \quad \square
\end{aligned}$$

- Discrete Case: $\text{Cov}(Y_1, Y_2) = \sum_{y_1}\sum_{y_2} y_1 y_2 p(y_1, y_2) - \mu_1\mu_2$
- Continuous Case: $\text{Cov}(Y_1, Y_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_1 y_2 f(y_1, y_2) dy_1 dy_2 - \mu_1\mu_2$

Example 1: Consider X, Y with the following joint PMF $p(x, y)$:

		y		
	$X\backslash Y$	0	1	2
x	0	1/8	1/8	0
	1	1/8	2/8	1/8
	2	0	1/8	1/8

Compute $\text{Cov}(X, Y)$.

Covariance

Example 1: Consider X, Y with the following joint PMF $p(x, y)$:

	$X \backslash Y$	y: 0	1	2	$p_X(x)$	$xp_X(x)$
x	0	1/8	1/8	0	2/8	0
	1	1/8	2/8	1/8	4/8	1/2
	2	0	1/8	1/8	2/8	1/2
	$p_Y(y)$	2/8	4/8	2/8		
	$yp_Y(y)$	0	1/2	1/2		

Compute $\text{Cov}(X, Y)$.

Solution:

- $\text{Cov}(X, Y) = \sum_x \sum_y xyp(x, y) - \{\sum_x xp_X(x)\}\left\{\sum_y yp_Y(y)\right\}$ Formula
  - $\sum_x xp_X(x) = 0 + 1/2 + 1/2 = 1$
  - $\sum_y yp_Y(y) = 0 + 1/2 + 1/2 = 1$
  - $\sum_x \sum_y xyp(x, y) = (0)(0)(1/8) + (0)(1)(1/8) + (0)(2)(0) + (1)(0)(1/8) + (1)(1)(2/8) + (1)(2)(1/8) + (2)(0)(0) + (2)(1)(1/8) + (2)(2)(1/8) = 10/8$
- Therefore, $\text{Cov}(X, Y) = 10/8 - (1)(1) = 1/4$. □

Example 2: (Continuous)
Suppose X and Y are continuous random variables on the unit square $[0,1]\times[0,1]$ with joint density $f(x,y)=2x^3+2y^3$. Compute $\text{Cov}(X,Y)$.

Solution:

- Formula: $\text{Cov}(X,Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)dxdy-\left\{\int_{-\infty}^{\infty}xf_X(x)dx\right\}\left\{\int_{-\infty}^{\infty}yf_Y(y)dy\right\}$
  - $f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy=\int_0^1 2x^3+2y^3dy=(2x^3y+\frac{2}{4}y^4)\big|_0^1=2x^3+\frac{1}{2}$
  - $\int_{-\infty}^{\infty}xf_X(x)dx=\int_0^1 x\left(2x^3+\frac{1}{2}\right)dx=\int_0^1 2x^4+\frac{1}{2}xdx=(\frac{2}{5}x^5+\frac{1}{4}x^2)\big|_0^1=\frac{13}{20}$
  - $f_Y(y)=\int_{-\infty}^{\infty}f(x,y)dx=\int_0^1 2x^3+2y^3dx=(\frac{2}{4}x^4+2y^3x)\big|_0^1=\frac{1}{2}+2y^3$
  - $\int_{-\infty}^{\infty}yf_Y(y)dy=\int_0^1 y\left(\frac{1}{2}+2y^3\right)dy=\int_0^1\frac{1}{2}y+2y^4dy=(\frac{1}{4}y^2+\frac{2}{5}y^5)\big|_0^1=\frac{13}{20}$
  - $$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)dxdy&=\int_0^1\int_0^1 xy\left(2x^3+2y^3\right)dxdy\\&=\int_0^1\int_0^1 2x^4y+2xy^4dxdy\\&=\int_0^1\left(\frac{2}{5}x^5y+x^2y^4\right)\Big|_0^1dy\\&=\int_0^1\frac{2}{5}y+y^4dy=\left(\frac{1}{5}y^2+\frac{1}{5}y^5\right)\Big|_0^1=\frac{2}{5}\end{aligned}$$
- Thus, $\text{Cov}(X,Y)=\frac{2}{5}-(\frac{13}{20})(\frac{13}{20})=-\frac{9}{400}$. □

Covariance

Properties: Covariance of Independent RVs

2. If Y_1 and Y_2 are independent random variables, then

$$\text{Cov}(Y_1, Y_2) = 0.$$

WARNING: The converse is false. Zero covariance DOES NOT always imply independence.

Proof:

$$\begin{aligned}
\text{Cov}(Y_1, Y_2) &= E(Y_1 Y_2) - \mu_1\mu_2 && \text{more useful formula for covariance} \\
&= E(Y_1)E(Y_2) - \mu_1\mu_2 && \text{expected value of independent RVs} \\
&= \mu_1\mu_2 - \mu_1\mu_2 && \text{def'n of expected value} \\
&= 0.
\end{aligned}$$

□

Covariance

Example 3: (Zero covariance does not imply independence)

Let X be a random variable that takes values $-2, -1, 0, 1, 2$, each with probability $1/5$. Let $Y = X^2$. Show that $\text{Cov}(X, Y) = 0$ but X and Y are not independent.

Solution:

- Computing $\text{Cov}(X, Y)$: Formula: $\text{Cov}(Y_1, Y_2) = \sum_{y_1} \sum_{y_2} y_1 y_2 p(y_1, y_2) - \mu_1 \mu_2$
  - The joint probability table of X and Y is as follows:

		x						
	$Y \backslash X$	-2	-1	0	1	2	$p_Y(y)$	$yp_Y(y)$
	0	0	0	1/5	0	0	1/5	0
y	1	0	1/5	0	1/5	0	2/5	2/5
	4	1/5	0	0	0	1/5	2/5	8/5
	$p_X(x)$	1/5	1/5	1/5	1/5	1/5		
	$xp_X(x)$	-2/5	-1/5	0	1/5	2/5		

  - $\mu_x = \sum_x xp_X(x) = -2/5 - 1/5 + 0 + 1/5 + 2/5 = 0$
  - $\mu_y = \sum_y yp_Y(y) = 0 + 2/5 + 8/5 = 2$
  - $\sum_x \sum_y xyp(x, y) = 0$
- Thus, $\text{Cov}(X, Y) = 0 - (0)(2) = 0$.
- Also, X and Y are clearly dependent because knowing the value X will give us the value of Y since $Y = X^2$. □

Properties: Variance is the Covariance of a RV w/ Itself

3. $\text{Cov}(Y_1, Y_1) = V(Y_1)$.

Proof:

$$\begin{aligned} \text{Cov}(Y_1, Y_1) &= E\{(Y_1 - \mu_1)(Y_1 - \mu_1)\} && \text{def'n of covariance} \\ &= E\{(Y_1 - \mu_1)^2\} \\ &= V(Y_1). && \text{def'n of variance} \end{aligned}$$

We can also use the more useful formula for covariance to do the proof...

$$\begin{aligned} \text{Cov}(Y_1, Y_1) &= E(Y_1 Y_1) - \mu_1\mu_1 && \text{more useful formula for covariance} \\ &= E(Y_1^2) - \mu_1^2 \\ &= V(Y_1). && \text{more useful formula for variance} \end{aligned}$$

□

- This property tells us that the covariance is a more general version of variance...

Covariance

Properties: Covariance of Linear Transformations

4. $\text{Cov}(aY_1 + b, cY_2 + d) = ac\text{Cov}(Y_1, Y_2)$,
   for constants a, b, c, and d.

Proof:

$$
\begin{aligned}
\text{Cov}(aY_1 + b, cY_2 + d) &= E\{(aY_1 + b)(cY_2 + d)\} - E(aY_1 + b)E(cY_2 + d) \\
&\quad \text{more useful formula for covariance: } \text{Cov}(Y_1, Y_2) = E(Y_1 Y_2) - \mu_1 \mu_2 \\
&= E(acY_1 Y_2 + bcY_2 + adY_1 + bd) \\
&\quad -\{aE(Y_1) + b\}\{cE(Y_2) + d\} \\
&= acE(Y_1 Y_2) + bcE(Y_2) + adE(Y_1) + bd \\
&\quad -acE(Y_1)E(Y_2) - bcE(Y_2) - adE(Y_1) - bd \\
&\quad \text{linearity of expectation} \\
&= acE(Y_1 Y_2) - acE(Y_1)E(Y_2) \quad \text{cancel out terms} \\
&= ac\{E(Y_1 Y_2) - E(Y_1)E(Y_2)\} \quad \text{factor out constants} \\
&= ac\text{Cov}(Y_1, Y_2). \quad \text{def'n of covariance} \quad \square
\end{aligned}
$$

Properties: Covariance of Sums

5. $\text{Cov}(Y_1 + Y_2, Y_3) = \text{Cov}(Y_1, Y_3) + \text{Cov}(Y_2, Y_3)$

Proof:

$$\begin{aligned}
\text{Cov}(Y_1 + Y_2, Y_3) &= E\{(Y_1 + Y_2)Y_3\} - E(Y_1 + Y_2)E(Y_3) \\
&\quad \text{more useful formula for covariance: } \text{Cov}(Y_1, Y_2) = E(Y_1 Y_2) - \mu_1\mu_2 \\
&= E(Y_1 Y_3 + Y_2 Y_3) - \{E(Y_1) + E(Y_2)\}E(Y_3) \\
&= E(Y_1 Y_3) + E(Y_2 Y_3) - E(Y_1)E(Y_3) - E(Y_2)E(Y_3) \\
&\quad \text{linearity of expectation} \\
&= \{E(Y_1 Y_3) - E(Y_1)E(Y_3)\} + \{E(Y_2 Y_3) - E(Y_2)E(Y_3)\} \\
&\quad \text{rearrange and group terms} \\
&= \text{Cov}(Y_1, Y_3) + \text{Cov}(Y_2, Y_3). \quad \text{def'n of covariance}
\end{aligned}$$

□

Covariance

Properties: Variance of the Sum

6. $V(Y_1 + Y_2) = V(Y_1) + V(Y_2) + 2\text{Cov}(Y_1, Y_2)$

Proof:

$$
\begin{aligned}
V(Y_1 + Y_2) &= E\{(Y_1 + Y_2)^2\} - \{E(Y_1 + Y_2)\}^2 \\
&\quad \text{more useful formula for variance: } V(Y) = E(Y^2) - \{E(Y)\}^2 \\
&= E(Y_1^2 + 2Y_1Y_2 + Y_2^2) - \{E(Y_1) + E(Y_2)\}^2 \\
&= E(Y_1^2) + 2E(Y_1Y_2) + E(Y_2^2) - \{E(Y_1)\}^2 - 2E(Y_1)E(Y_2) - \{E(Y_2)\}^2 \\
&\quad \text{linearity of expectation} \\
&= \{E(Y_1^2) - \{E(Y_1)\}^2\} + \{E(Y_2^2) - \{E(Y_2)\}^2\} + 2E(Y_1Y_2) - 2E(Y_1)E(Y_2) \\
&\quad \text{rearrange and group terms from the same RV} \\
&= V(Y_1) + V(Y_2) + 2\{E(Y_1Y_2) - E(Y_1)E(Y_2)\} \quad \text{def'n of variance} \\
&= V(Y_1) + V(Y_2) + 2\text{Cov}(Y_1, Y_2). \quad \text{def'n of covariance} \quad \square
\end{aligned}
$$

- This property also implies: $V(Y_1 - Y_2) = V(Y_1) + V(Y_2) - 2\text{Cov}(Y_1, Y_2)$.
- If Y_1, Y_2 are independent: $V(Y_1 + Y_2) = V(Y_1) + V(Y_2)$. independent $\Rightarrow \text{Cov}(Y_1, Y_2) = 0$

Correlation

Correlation

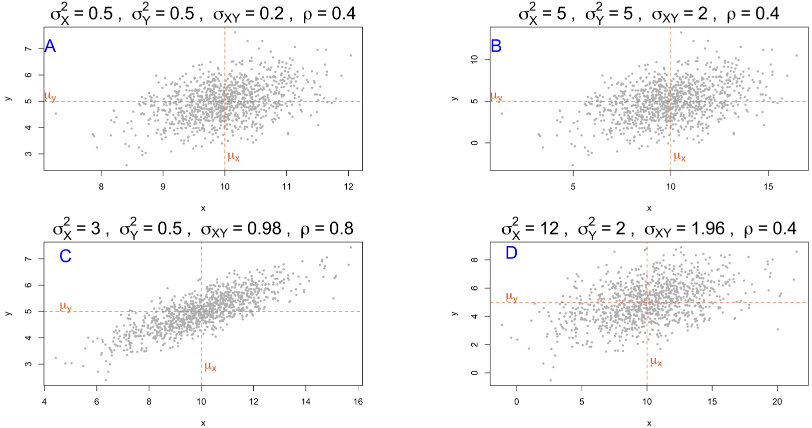
Definition: Correlation

If Y_1 and Y_2 are random variables with finite second moments, the *correlation* of Y_1 and Y_2 is

$$\text{Corr}(Y_1, Y_2) = \frac{\text{Cov}(Y_1, Y_2)}{\sqrt{V(Y_1)V(Y_2)}}.$$

- Notation: $\rho_{12} = \text{Corr}(Y_1, Y_2)$ The Greek letter ρ, pronounced as 'rho', is reserved for correlation.
- The units of covariance $\text{Cov}(Y_1, Y_2)$ are 'units of Y_1 times units of Y_2'. This makes it hard to compare covariances...
- Correlation is a way to remove the scale from the covariance.
- Correlation takes the covariance and makes it meaningful.

Correlation takes the covariance and makes it meaningful...



Which X and Y pair has the strongest linear relationship?

- The correlation in C is closer to 1 than in A, B, and D.
- We can easily increase the covariance by increasing the variance w/o changing the correlation: $\rho = \frac{\sigma_{XY}}{\sqrt{\sigma_X^2 \sigma_Y^2}} \Rightarrow \sigma_{XY} = \rho\sqrt{\sigma_X^2 \sigma_Y^2}$.

Correlation vs. Covariance

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}$$

Covariance	**Correlation**
indicates the nature of the relationship between X and Y (positive or negative)	measure of the strength of the relationship between X and Y
magnitude of the covariance is affected by the magnitude of X and Y	independent of the influence of the magnitude of X and Y
not standardized	standardized
Values can range between $-\infty$ and ∞	Values can range between -1 and 1
Exact value of the covariance is not useful but the sign indicates the relationship between X and Y.	The closer the value is to -1 or 1, the stronger the relationship between X and Y.

Properties: Correlation always falls between -1 and 1

1. $-1 \leq \text{Corr}(Y_1, Y_2) \leq 1.$

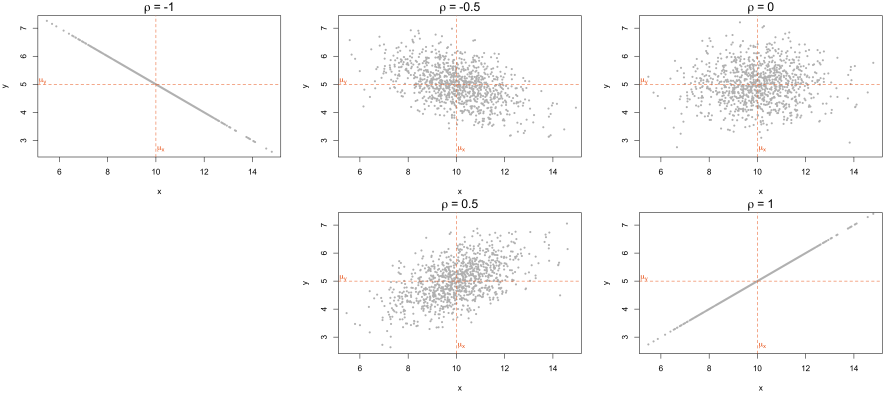
Proof:

$$\begin{aligned}
\{\text{Cov}(Y_1, Y_2)\}^2 &\leq V(Y_1)V(Y_2) && \text{Cauchy-Schwarz inequality} \\
\Rightarrow -\sqrt{V(Y_1)V(Y_2)} &\leq \text{Cov}(Y_1, Y_2) \leq \sqrt{V(Y_1)V(Y_2)} \\
\Rightarrow -1 &\leq \frac{\text{Cov}(Y_1, Y_2)}{\sqrt{V(Y_1)V(Y_2)}} \leq 1 && \text{multiply } \frac{1}{\sqrt{V(Y_1)V(Y_2)}} \\
\Rightarrow -1 &\leq \text{Corr}(Y_1, Y_2) \leq 1 && \text{def'n of correlation}
\end{aligned}$$

□

Correlation

- The closer ρ_{XY} is to 1 or -1, the stronger the linear relationship.
- If $\rho_{XY} = -1$, X and Y has a perfect negative linear relationship.
- If $\rho_{XY} = 1$, X and Y has a perfect positive linear relationship.



Properties: A RV is Perfectly Correlated with Itself

❷ $\text{Corr}(Y_1, Y_1) = 1.$

Proof:

$$\begin{aligned}\text{Corr}(Y_1, Y_1) &= \frac{\text{Cov}(Y_1, Y_1)}{\sqrt{V(Y_1)V(Y_1)}} && \text{def'n of covariance} \\ &= \frac{\text{Cov}(Y_1, Y_1)}{V(Y_1)} \\ &= \frac{V(Y_1)}{V(Y_1)} && \text{Property \# 3: } \text{Cov}(Y_1, Y_1) = V(Y_1) \\ &= 1.\end{aligned}$$

□

Example 4: (Recall Ex. 1)

Consider X, Y with the following joint PMF $p(x, y)$:

	$X \backslash Y$	y: 0	1	2
	0	1/8	1/8	0
x	1	1/8	2/8	1/8
	2	0	1/8	1/8

Compute Corr(X, Y).

Correlation

Example 4: (Recall Ex. 1)

Consider X, Y with the following joint PMF $p(x, y)$:

	$X \backslash Y$	y: 0	1	2	$p_X(x)$	$x^2 p_X(x)$
x	0	1/8	1/8	0	2/8	0
	1	1/8	2/8	1/8	4/8	1/2
	2	0	1/8	1/8	2/8	1
	$p_Y(y)$	2/8	4/8	2/8		
	$y^2 p_Y(y)$	0	1/2	1		

Compute $\text{Corr}(X, Y)$.

Solution:

- Formula: $\text{Corr}(X, Y) = \frac{\text{Cov}(X,Y)}{\sqrt{V(X)V(Y)}}$
  - Earlier we computed $\text{Cov}(X, Y) = 1/4$.
  - $V(X) = E(X^2) - \{E(X)\}^2$
  - Earlier we computed $E(X) = 1$.
  - $E(X^2) = \sum_x x^2 p_X(x) = 0 + 1/2 + 1 = \frac{3}{2}$
  - $V(X) = \frac{3}{2} - 1^2 = \frac{1}{2}$

(cont'd next slide...)

Correlation

Example 4: (Recall Ex. 1)

Consider X, Y with the following joint PMF $p(x, y)$:

	$X \backslash Y$	y = 0	1	2	$p_X(x)$	$x^2 p_X(x)$
x	0	1/8	1/8	0	2/8	0
	1	1/8	2/8	1/8	4/8	1/2
	2	0	1/8	1/8	2/8	1
	$p_Y(y)$	2/8	4/8	2/8		
	$y^2 p_Y(y)$	0	1/2	1		

Compute $\text{Corr}(X, Y)$.

Solution:

- Formula: $\text{Corr}(X, Y) = \frac{\text{Cov}(X,Y)}{\sqrt{V(X)V(Y)}}$
  - Earlier we computed $\text{Cov}(X, Y) = 1/4$.
  - $V(Y) = E(Y^2) - \{E(Y)\}^2$
  - Earlier we computed $E(Y) = 1$.
  - $E(Y^2) = \sum_y y^2 p_Y(y) = 0 + 1/2 + 1 = \frac{3}{2}$
  - $V(Y) = \frac{3}{2} - 1^2 = \frac{1}{2}$

(cont'd next slide...)

Example 4: (Recall Ex. 1)

Consider X, Y with the following joint PMF $p(x, y)$:

	$X \backslash Y$	y = 0	y = 1	y = 2	$p_X(x)$	$x^2 p_X(x)$
x	0	1/8	1/8	0	2/8	0
x	1	1/8	2/8	1/8	4/8	1/2
x	2	0	1/8	1/8	2/8	1
	$p_Y(y)$	2/8	4/8	2/8		
	$y^2 p_Y(y)$	0	1/2	1		

Compute $\text{Corr}(X, Y)$.

Solution:

- Formula: $\text{Corr}(X, Y) = \frac{\text{Cov}(X,Y)}{\sqrt{V(X)V(Y)}}$
    - Earlier we computed $\text{Cov}(X, Y) = 1/4$.
    - $V(X) = \frac{3}{2} - 1^2 = \frac{1}{2}$
    - $V(Y) = \frac{3}{2} - 1^2 = \frac{1}{2}$
- Thus, $\text{Corr}(X, Y) = \frac{\frac{1}{4}}{\sqrt{\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}} = \frac{1}{2}$. □

Covariance

Example 5: (Recall Ex. 2)
Suppose X and Y are continuous random variables on the unit square $[0,1] \times [0,1]$ with joint density $f(x,y) = 2x^3 + 2y^3$. Compute $\text{Corr}(X,Y)$.

Solution:

- Formula: $\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{V(X)V(Y)}}$
  - Earlier we computed $\text{Cov}(X,Y) = -\frac{9}{400}$.
  - $V(X) = E(X^2) - \{E(X)\}^2$
  - Earlier we computed $E(X) = \frac{13}{20}$ and $f_X(x) = 2x^3 + \frac{1}{2}$.
  - $E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^1 x^2 \left(2x^3 + \frac{1}{2}\right) dx = \int_0^1 2x^5 + \frac{1}{2}x^2 dx = \left(\frac{1}{3}x^6 + \frac{1}{6}x^3\right)\Big|_0^1 = \frac{1}{2}$
  - $V(X) = \frac{1}{2} - \left(\frac{13}{20}\right)^2 = \frac{31}{400}$

(cont'd next slide...)

Covariance

Example 5: (Recall Ex. 2)
Suppose X and Y are continuous random variables on the unit square $[0,1] \times [0,1]$ with joint density $f(x,y) = 2x^3 + 2y^3$. Compute $\text{Corr}(X,Y)$.

Solution:

- Formula: $\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{V(X)V(Y)}}$
  - Earlier we computed $\text{Cov}(X,Y) = -\frac{9}{400}$.
  - $V(X) = \frac{1}{2} - \left(\frac{13}{20}\right)^2 = \frac{31}{400}$
  - $V(Y) = E(Y^2) - \{E(Y)\}^2$
  - Earlier we computed $E(Y) = \frac{13}{20}$ and $f_Y(y) = \frac{1}{2} + 2y^3$.
  - $E(Y^2) = \int_{-\infty}^{\infty} y^2 f_Y(y)dy = \int_0^1 y^2\left(\frac{1}{2} + 2y^3\right)dy = \int_0^1 \frac{1}{2}y^2 + 2y^5 dy = \left(\frac{1}{6}y^3 + \frac{1}{3}y^6\right)\Big|_0^1 = \frac{1}{2}$
  - $V(Y) = \frac{1}{2} - \left(\frac{13}{20}\right)^2 = \frac{31}{400}$
- Thus, $\text{Corr}(X,Y) = \frac{-\frac{9}{400}}{\sqrt{\left(\frac{31}{400}\right)\left(\frac{31}{400}\right)}} = -0.29.$ □

Conditional Expectation

Conditional Expectation

Definition: Conditional Expectation

If Y_1 and Y_2 are any two random variables, the *conditional expectation* of $g(Y_1)$, given that $Y_2 = y_2$, is

$$E\{g(Y_1)|Y_2 = y_2\} = \int_{-\infty}^{\infty} g(y_1)f(y_1|y_2)dy_1,$$

if Y_1 and Y_2 are jointly continuous and

$$E\{g(Y_1)|Y_2 = y_2\} = \sum_{\text{all } y_1} g(y_1)p(y_1|y_2),$$

if Y_1 and Y_2 are jointly discrete.

- If $g(Y_1) = Y_1$, $E\{g(Y_1)|Y_2 = y_2\}$ is the conditional expectation.
- If $g(Y_1) = (Y_1 - \mu_1)^2$, $E\{g(Y_1)|Y_2 = y_2\}$ is the conditional variance.

Conditional PMF

Example 6: (Recall Ex. 5 in Lec 16)

Suppose X and Y have the following joint PMF:

	$X \backslash Y$	y: 0	1	2
	0	0.10	0.04	0.02
x	1	0.08	0.20	0.06
	2	0.06	0.14	0.30

Find $E(Y|X=1)$.

Solution:

$$\begin{aligned} E(Y|X=1) &= 0 \times P(Y=0|X=1) + 1 \times P(Y=1|X=1) + 2 \times P(Y=2|X=1) \\ &\quad \text{def'n of conditional expectation} \\ &= 0 \times 0.2353 + 1 \times 0.5882 + 2 \times 0.1765 \\ &\quad \text{We solved for the conditional PMF of } Y \text{ in Lec 16, Slide 31: } p(y|X=1) = \begin{cases} 0.2353, & \text{if } y=0 \\ 0.5882, & \text{if } y=1 \\ 0.1765, & \text{if } y=2. \end{cases} \\ &= 0.9412. \end{aligned}$$

□

Conditional Expectation

Properties: Conditional Expectation

1. *Law of Total Expectation:* Let Y_1 and Y_2 be any two random variables. Then,

$$E(Y_1) = E\{E(Y_1|Y_2)\}.$$

Remarks:

- $E(Y_1|Y_2)$ is a random variable and we can compute its expectation.
- Sometimes called the "Law of Iterated Expectations"...
- It is a decomposition rule.
- It decomposes $E(Y_1)$ into smaller/ easier conditional expectations.

A Closer Look At the Law of Total Expectation...

Suppose we have X and Y continuous RVs. Then, by the law of total expectation, we have: $E(X) = E\{E(X|Y)\}$.

WHY DOES THIS WORK?

- We know $E(X)$ is a constant/number and is simply

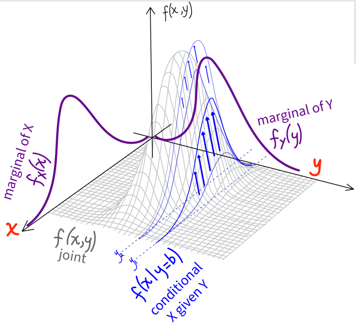
$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx. \quad \text{def'n of expected value}$$

- Now, apply the def'n of expected value to $E\{E(X|Y)\}$:

$$E\{E(X|Y)\} = \int_{-\infty}^{\infty} E(X|Y=y) f_Y(y) dy \quad \text{WHY?}$$

(cont'd next slide...)

Conditional Expectation



Claim: $E(X|Y)$ is a random variable.

- What are the possible values of $E(X|Y)$?
  - $E(X|Y = y_1)$
  - $E(X|Y = y_2)$

    $\vdots$
  - $E(X|Y = y_n)$
- The value of $E(X|Y)$ depends on the value of the random variable Y.
- In fact, $E(X|Y)$ is a <u>function of the random variable Y</u>.
  Mathematically, we write this as: $g(Y) = E(X|Y)$.

How to find the expectation of a function of the random variable Y?

Recall Theorem 4.4 in Lec 9, Slide 20:
Let $g(Y)$ be a function of Y. Then the expected value of $g(Y)$ is given by

$$E\{g(Y)\} = \int_{-\infty}^{\infty} g(y) f_Y(y) dy,$$

provided that the integral exists.

(cont'd next slide...)

Conditional Expectation

A Closer Look At the Law of Total Expectation... (cont'd)

$$
\begin{aligned}
E\{E(X|Y)\} &= \int_{-\infty}^{\infty} E(X|Y=y) f_Y(y)dy && \text{prev. slide: } E\{g(Y)\} = \int_{-\infty}^{\infty} g(y)f_Y(y)dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\underline{f(x|y)}dx\underline{f_Y(y)}dy \\
& && \text{def'n of conditional expectation: } E\{g(Y_1)|Y_2=y_2\} = \int_{-\infty}^{\infty} g(y_1)f(y_1|y_2)dy_1 \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\underline{f(x,y)}dxdy && \text{def'n of conditional PDF: } f(y_1|y_2) = \frac{f(y_1,y_2)}{f_2(y_2)} \\
&= \int_{-\infty}^{\infty} x\left\{\int_{-\infty}^{\infty} f(x,y)dy\right\}dx && \text{rearranging} \\
&= \int_{-\infty}^{\infty} x f_X(x)dx && \text{def'n of marginal PDF: } f_1(y_1) = \int_{-\infty}^{\infty} f(y_1,y_2)dy_2 \\
&= E(X). && \text{def'n of expected value}
\end{aligned}
$$

Same principle applies for discrete RV. □

Properties: Conditional Expectation

❷ *Law of Total Variance:* Let Y_1 and Y_2 be any two random variables. Then,

$$V(Y_1) = E\{V(Y_1|Y_2)\} + V\{E(Y_1|Y_2)\}.$$

Proof:

$$\begin{aligned}
V(Y_1) &= E(Y_1^2) - \{E(Y_1)\}^2 \quad \text{def'n of variance: } V(X) = E(X^2) - \{E(X)\}^2 \\
&= E\{E(Y_1^2|Y_2)\} - [E\{E(Y_1|Y_2)\}]^2 \\
&\quad \text{law of total expectation: } E(Y_1) = E\{E(Y_1|Y_2)\} \\
&= E[V(Y_1|Y_2) + \{E(Y_1|Y_2)\}^2] - [E\{E(Y_1|Y_2)\}]^2 \\
&\quad \text{def'n of variance: } E(X^2) = V(X) + \{E(X)\}^2 \\
&= E\{V(Y_1|Y_2)\} + \underline{E[\{E(Y_1|Y_2)\}^2] - [E\{E(Y_1|Y_2)\}]^2} \\
&\quad \text{linearity of expectation} \\
&= E\{V(Y_1|Y_2)\} + \underline{V\{E(Y_1|Y_2)\}}. \\
&\quad \text{def'n of variance: } V(X) = E(X^2) - \{E(X)\}^2
\end{aligned}$$

Conditional Expectation

Example 7:

Let N be the number you get when you roll a die. Let H be the number of heads after tossing a fair coin N times. Find $E(H)$.

Solution:

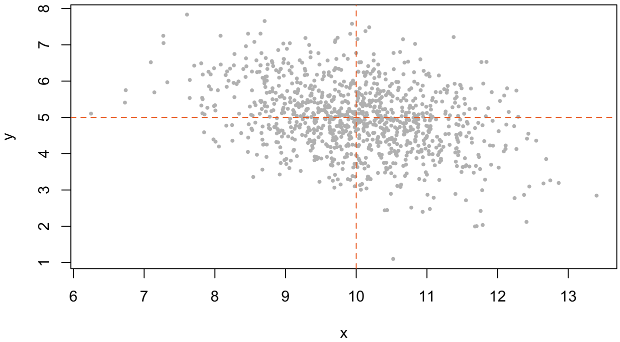
$$
\begin{aligned}
E(H) &= E\{E(H|N)\} \quad \text{law of total expectation} \\
& \quad \text{Note: } H \text{ is a Binomial RV with prob. of success} = \frac{1}{2} \text{ and } N \text{ num. of trials} \\
&= E\left(\frac{N}{2}\right) \quad \text{If } X \sim B(n,p),\ E(X) = np. \text{ Here, } n = N \text{ and } p = \frac{1}{2}. \\
&= \frac{1}{2}E(N) \quad \text{linearity of expectation} \\
&= \frac{1}{2}\left(1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6}\right) \\
& \quad \text{def'n of expected value} \\
&= \frac{3.5}{2} = 1.75.
\end{aligned}
$$

□

Bonus Exercise

Bonus Exercise

Guess the correlation of X and Y...



- If you guess the EXACT number: +10 pts
- If you are close enough: +5 pts

Email your guess to: marylai.salvana@uconn.edu

Questions?

Homework Exercises: 5.21, 5.23, 5.27, 5.35, 5.41

Solutions will be discussed this Friday by the TA.