Stat 401
Fall 2017

Shyamala Nagaraj

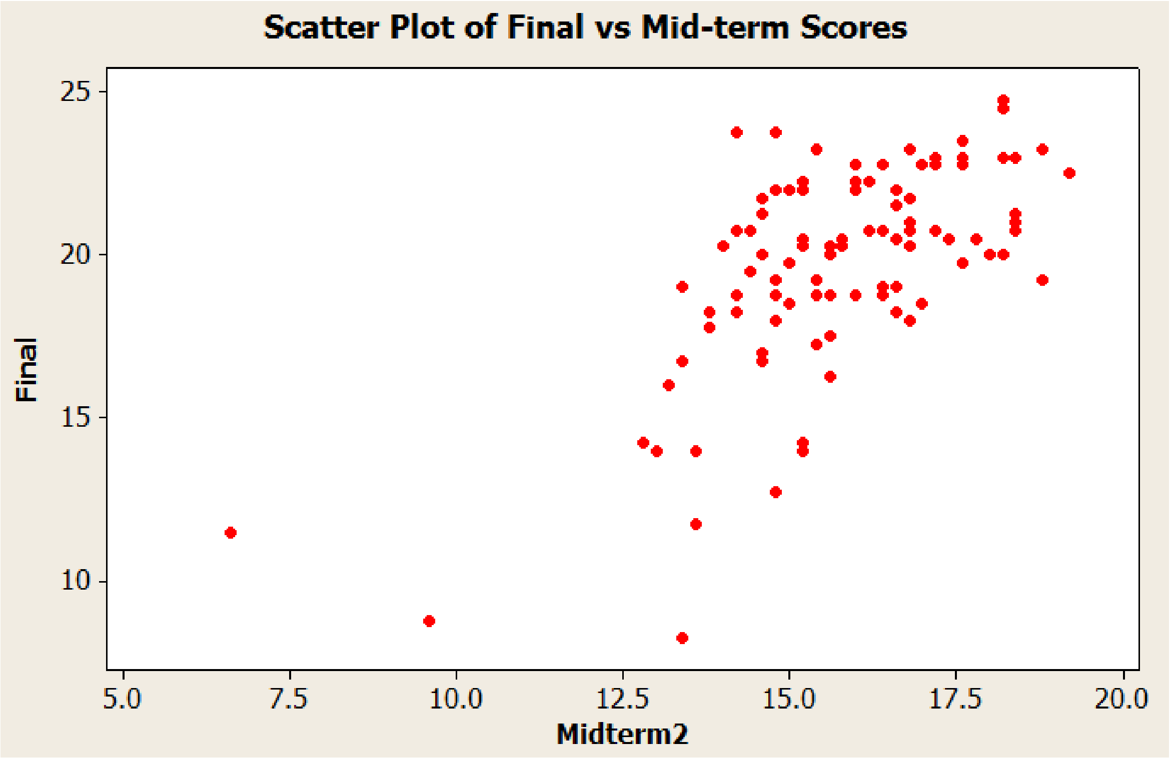
Vijay Nair

Slide Set #3

Bivariate Random Variables, Topics:

- Motivation and Notation
- Joint Distributions of Discrete Random Variables
  - Joint and Marginal Distributions (sample to population)
  - Expectation and Variance
  - Covariance and Correlation
  - Independence
- Analogous Concepts for Continuous Distributions
- Linear Combination of Random Variables
  - Expectation
  - Variance

Recall this example from Slides Set 1



<u>Bivariate Characteristics</u>

Relative Frequency Distributions

Summary Statistics:

Correlation = 0.66 ← what is this? Is this a good summary?

What are analogous concepts for populations?

Bivariate Random Variables

- A pair of random variables: Notation (X, Y)
- Sample data (y_i, x_i), i=1, …n
- More examples:
  - Heights of students in Pioneer High School and their fathers
  - Number of wins of a sample of women's college basketball teams in 2014 and 2015
  - Amount of snow in December 2014 and 2105 for a sample of cities in the US
  - BP measurements of selected patients: (systolic BP, diastolic BP)
  - Incomes of spouses: (Income F, Income M)

Notation for Populations

- Random Variable X, Y, Z, etc.
  - Expected Value, also called Mean: μ_X, μ_Y, μ_Z, etc.
  - Variance: $\sigma_X^2, \sigma_Y^2, \sigma_Z^2$, etc.
  - Standard Deviation (or SD): $\sigma_X, \sigma_Y, \sigma_Z$, etc.

- Random Variable X_1, X_2, etc.
  - Expected Value or Mean: μ_1, μ_2, etc.
  - Variance σ_1^2, σ_2^2, etc.
  - Standard Deviation (or SD): σ_1, σ_2, etc.

- Problems of interest: Joint distribution of X and Y; relationship between X and Y
  - Study their distribution of X and Y together and not just separately
  - So … what is "Joint Distribution"? All possible values (x,y) and their probabilities

Bivariate Distribution for Discrete RVs -- Population

- Randomly selected UM undergraduate
- X = Dyes hair (yes/no)
  - Possible values (simplified): {0, 1}
- Y = Pierces ear (none, one ear, both ears)
  - Possible values (simplified): {0, 1, 2}
- All possible values of (X, Y)? How many?
- Joint distribution of (X, Y) given below:

Possible values y

Possible values x		0	1	2	Marginal Dist of X
	0	0.4	0.1	0	0.5
	1	0.1	0.3	0.1	0.5
	Marginal Dist of Y	0.5	0.4	0.1	Total 1.0

What are some of the questions of interest?
How is the joint distribution useful?

Bivariate Distribution for Discrete RVs: A Closer Look

- Joint pmf of (X, Y): all possible values (x, y) and their probabilities
- Probability in each cell: P(X = x, Y = y) → notation: p(x, y)

  → interpretation: P(X = x and Y = y)
- Sum of all values = 1
- Marginal pmf of X → sum of probabilities across columns: $p_X(x)$
- Marginal pmf of Y → sum of probabilities across rows: $p_Y(y)$
- **Conditional pmf of Y given X=0→ sum of probabilities across row X=0 divided by row probability: $p_Y(y|X=0)$**
- What is the probability that a randomly selected student has dyed hair but no pierced ears?
- What is the probability that among those with dyed hair, a randomly selected student has no pierced ears?

Expected value and variance of h(X, Y)

- Let h(X, Y) be a new random variable that is a function of (depends on) the bivariate random variable (X, Y)
- Special case: h(X, Y) = X or h(X, Y) = Y
- Can use the marginal pmf of X or Y to compute expected value and variance
- Another special case: h(X, Y) = aX + bY + c → linear combination
- More complex : h(X, Y) = XY
- E[h(X, Y)] = sum of [h(x, y) times p(x, y)]over all possible values of (x, y)
- var[h(X, Y)] = sum of [h(x, y) – E[h(X,Y)]^2 times p(x, y)]over all possible values of (x, y)

Example

- There are three fixed-price dinners at a restaurant: \$12, \$15, \$20. Suppose you pick a randomly selected pair (of man and woman) having dinner at the restaurant:
- Let X = cost of the man's dinner; Y = cost of woman's dinner

– What are the marginal pmf's of X and Y?

– Is E(X) larger than E(Y)?

– What is the probability that the man's dinner costs more than the woman's?

$p(x, y)$		y = 12	15	20
x	12	.05	.05	.10
	15	.05	.10	.35
	20	0	.20	.10

- Suppose the customers are told after the dinner that they will get a refund of the difference between the costs of the more expensive and less expensive meal.

– What is h(X, Y)?

– What is the expected refund?

– If customers know this before hand and want to maximize the refund, what will the joint distribution be?

Bivariate RV (X, Y): Covariance of X and Y

- Covariance is of special interest
- Cov(X, Y) = E [(X – μ_X)(Y – μ_Y)] ← definition
  - Variance can be thought of as a special case
  - Sample version?
  - Interpretation – **covarying**
  - Interpretations when negative, positive, or (close to) zero
- Computation:
  - Take each possible value of (x – μ_X)(y- μ_Y)
  - Multiply by its probability
  - Sum up over all values to get covariance
- X = Dyes hair; Y = Pierces ear

	0	**1**	**2**
0	**0.4**	**0.1**	**0**
1	**0.1**	**0.3**	**0.1**

→

	-0.6	**0.4**	**1.4**
- 0.5	**0.4**	**0.1**	**0**
+ 0.5	**0.1**	**0.3**	**0.1**

- **Cov(X, Y) = E [(X – μ_X)(Y – μ_Y)] = sum of (x – μ_X)(y- μ_Y) times the probabilities**
  = (-0.5)(-0.6 (0.4) + (-0.5)(0.4)(0.1) + … + (0.5)(1.4)(0.1) = ?

Covariance of X and Y: Alternate expression

- cov(X, Y) = E [(X – μ_X)(Y – μ_Y)] → notation σ_{XY}
- E [(X – μ_X)(Y – μ_Y)] = E(XY) – (μ_X μ_Y) → Why?
- Easier to compute
- E(XY) = take each possible value of (xy), multiply by its probability and sum over all values
- X = Dyes hair; Y = Pierces ear

	0	**1**	**2**
0	0.4	0.1	0
1	0.1	0.3	0.1

- E(XY) = (0)(0) (0.4) + (0) (1) (0.1) + (0)(2)(0) + (1)(0)(0.1) + (1)(1)(0.3) + (1)(2)(0.1)
  = 0.3 + 0.2 = 0.5
- μ_X = 0.5; μ_y = 0.6
- cov(X, Y) or σ_{XY} = E(XY) – (μ_X μ_Y) = 0.5 – (0.5)(0.6) = 0.5 – 0.3 = 0.2 ← positive
- Positive relationship between dying hair and piercing ear
- **Is this a strong relationship?**

Correlation (and Covariance)

- Cov(X, Y) → hard to interpret the magnitude
- Depends on the variability of X and Y (units of observations, etc.)

Standardize the covariance → divide by the product of the standard deviations

$$Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

- Correlation is dimension free (no unit of measurement)
- Value is always between – 1 and 1
- Correlation of – 1 → strong negative (linear) relationship between X and Y
- Correlation of + 1 → strong positive (linear) relationship between X and Y
- Correlation of 0 → no linear relationship between X and Y

- **Correlation is not the same thing as independence**

Back to hair/ear example

- X = dyes hair; Y = pierces ear

	0	**1**	**2**
0	0.4	0.1	0
1	0.1	0.3	0.1

- cov(X, Y) = E(XY) – ($\mu_X \mu_Y$) = 0.5 – (0.5)(0.6) = 0.5 – 0.3 = 0.2 (from before)

- var(X) = E(X^2) – $(\mu_X)^2$ = (1^2)(0.5) – $(0.5)^2$ = 0.25 → σ_X = 0.5
- var(Y) = E(Y^2) – $(\mu_Y)^2$ = (1^2)(0.4) + (2^2)(0.1) – ((0.6)^2 = 0.44 → σ_Y = 0.663

- **Cor(X, Y) = 0.2/(0.0.5)(0.663) = 0.60**

- Let's consider the question from before:
- Positive relationship between dying hair and piercing ear
- **Is this a strong relationship?**

Independence

- Are X and Y independent for the joint pmf below?

	0	**1**	**2**	$p_X(x)$
0	0.25	0.2	0.05	0.5
1	0.25	0.2	0.05	0.5
$p_Y(Y) \rightarrow$	0.5	0.4	0.1	Total 1.0

- If X and Y are independent, you can get the joint pmf by taking the product of the two marginal pmfs

- $p(x, y) = p_X(x)\, p_Y(y)$ for all possible values (x, y)
- Same true for cdf: $F(x, y) = F_X(x)\, F_Y(y)$

- **If X and Y are independent, cov(X,Y) = 0 $\leftrightarrow$ cor(X, Y) = 0**
- **Why?**
- **Converse if not true: Cor(X, Y) = 0 does not mean they are independent**

Joint distribution for continuous RVs

- (X, Y) continuous RVs
- Joint distribution:
  - Joint pdf f(x, y) → non-negative; not a probability
  - Joint cdf F(x, y)

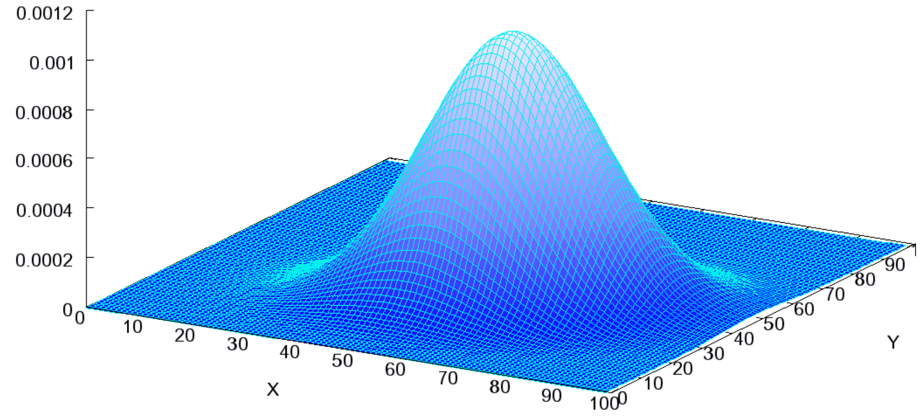
- Marginal pdf's of X and Y: $f_X(x)$ and $f_Y(y)$

- Computation of covariance and correlation based on the same ideas as before
- But now involve weighting by the pdf f(x, y) and integrating over the range of (x, y) rather than by weighting and summing over the possible values as in the discrete case

- Most interesting case for us:
  - Bivariate normal distribution of (X, Y)

Bivariate Normal Distribution

- Joint density of bivariate normal (from "Multivariate Gaussian". Licensed under CC BY-SA 3.0 via Commons - https://commons.wikimedia.org/wiki/File:Multivariate_Gaussian.png#/)



- Characterized by:
  - Mean and SD of X: μ_X and σ_X
  - Mean and SD of Y: μ_Y and σ_Y
  - Cor(X, Y) → sometimes called ρ – "rho" → *recall sample version is r*
  - Marginal distributions of X and Y are normal

Linear combination of bivariate RVs – expected value

- (X, Y) bivariate RV
- Discussed earlier: h(X, Y) → function of X and Y
- Special case: linear function or linear combination of RVs
  - h(X, Y) = aX + bY + c for some numbers a, b, and c
- **Examples:**
  - **X + Y, X – Y, 3X – 4Y + 10, (X + Y)/2**
  - **(X+Y)/2 = ½ X + ½ Y → sample mean**

$$\begin{aligned} E(aX + bY + c) &= E(aX) + E(bY) + E(c) \\ &= a\,E(X) + b\,E(Y) + c \\ &= a\,\mu_X + b\,\mu_Y + c \end{aligned}$$

- X and Y don't have to be independent.
- Examples above

Linear combination of many RVs – expected value

- Suppose we have n random variables: $X_1, X_2, \ldots, X_n$
- Let Y be a new RV given by the linear combination

$$Y = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$$

- Examples:
  - $Y = X_1 + X_2 + X_3$, or $Y = 2X_1 - 3X_2 + 10X_3 - 5X_4$
  - $Y = X_1 + X_2 + \ldots + X_n$ ← sum
  - $Y = (X_1 + X_2 + \ldots + X_n)/n$ ← sample mean

$$\begin{aligned} E(Y) &= E(a_1 X_1 + a_2 X_2 + \ldots + a_n X_n) \\ &= E(a_1 X_1) + E(a_2 X_2) + \ldots + E(a_n X_n) \\ &= a_1 \mu_1 + a_2 \mu_2 + \ldots + a_n \mu_n \end{aligned}$$

- The RVs do not have to be independent
- Application to examples

Linear combination of bivariate RVs – variance

- (X, Y) bivariate RV
- var(X + Y) = var(X) + var(Y) + 2 cov(X, Y)
- Intuition: take Y = X or – X
- Reason for the general result?

- What is var(X – Y)?

- **If X and Y are independent, var(X + Y) = var(X) + var(Y)**
- HOLDS ONLY WHEN X and Y ARE INDEPENDENT

- Suppose $X_1, X_2, \ldots, X_n$ are iid with mean μ and variance σ^2

- What is the expected value and the variance of the sample mean?

Linear combination of independent RVs – variance

- Suppose we have n random variables $X_1, X_2, \ldots, X_n$
- AND they are all independent
- Let Y be a new RV given by the linear combination

$$\mathbf{Y = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n}$$

- Examples:
  - $Y = X_1 + X_2 + X_3$, or $Y = 2X_1 - 3X_2 + 10X_3 - 5X_4$
  - $Y = X_1 + X_2 + \ldots + X_n$ ← sum
  - $Y = (X_1 + X_2 + \ldots + X_n)/n$ ← sample mean

$$\begin{aligned} Var(Y) &= Var(a_1 X_1 + a_2 X_2 + \ldots + a_n X_n) \\ &= Var(a_1 X_1) + Var(a_2 X_2) + \ldots + Var(a_n X_n) \\ &= a_1^2 Var(X_1) + a_2^2 Var(X_2) + \ldots + a_n^2 Var(X_n) \end{aligned}$$

- Example: what is the variance of the sum? the sample mean?
- Interpretations
- What is SD?