Random Variables and Probability Distributions

Math 130: Introductory Statistics

Definitions

- **Random variable**: a variable that has a single numerical value determined by chance.
- Data is a bunch of realizations of a random variable.
- **Discrete random variable**: an RV that can take on "countable" values.
- **Continuous random variable**: a RV that can take on infinitely many values on a continuous scale.

Definitions

- **Random variable**: a variable that has a single numerical value determined by chance.
- Data is a bunch of realizations of a random variable.
- **Discrete random variable**: an RV that can take on "countable" values.
- **Continuous random variable**: a RV that can take on infinitely many values on a continuous scale.

Definitions

- **Random variable**: a variable that has a single numerical value determined by chance.
- Data is a bunch of realizations of a random variable.
- **Discrete random variable**: an RV that can take on "countable" values.
- **Continuous random variable**: a RV that can take on infinitely many values on a continuous scale.

Definitions

- **Random variable**: a variable that has a single numerical value determined by chance.
- Data is a bunch of realizations of a random variable.
- **Discrete random variable**: an RV that can take on "countable" values.
- **Continuous random variable**: a RV that can take on infinitely many values on a continuous scale.

Probability distribution

- A **probability distribution** is a graph, table, or formula that gives the probability for each value of a random variable.
- The sum of the probabilities for all possible values an RV can take must equal 1 (or 100%).

$$\sum P(x_i) = 1$$

- Each individual probability must be between zero and one.

$$0 \leq P(x_i) \leq 1$$

Probability distribution

- A **probability distribution** is a graph, table, or formula that gives the probability for each value of a random variable.
- The sum of the probabilities for all possible values an RV can take must equal 1 (or 100%).

$$\sum P(x_i) = 1$$

- Each individual probability must be between zero and one.

$$0 \leq P(x_i) \leq 1$$

Probability distribution

- A **probability distribution** is a graph, table, or formula that gives the probability for each value of a random variable.
- The sum of the probabilities for all possible values an RV can take must equal 1 (or 100%).

$$\sum P(x_i) = 1$$

- Each individual probability must be between zero and one.

$$0 \leq P(x_i) \leq 1$$

Probability distribution

- A **probability distribution** is a graph, table, or formula that gives the probability for each value of a random variable.
- The sum of the probabilities for all possible values an RV can take must equal 1 (or 100%).

$$\sum P(x_i) = 1$$

- Each individual probability must be between zero and one.

$$0 \leq P(x_i) \leq 1$$

Probability distribution

- A **probability distribution** is a graph, table, or formula that gives the probability for each value of a random variable.
- The sum of the probabilities for all possible values an RV can take must equal 1 (or 100%).

$$\sum P(x_i) = 1$$

- Each individual probability must be between zero and one.

$$0 \leq P(x_i) \leq 1$$

Examples

- Let a random variable x be the number of children someone has.
- What kind of random variable is this?
- Lets construct a probability distribution for this class.

Examples

- Let a random variable x be the number of children someone has.
- What kind of random variable is this?
- Lets construct a probability distribution for this class.

Examples

- Let a random variable x be the number of children someone has.
- What kind of random variable is this?
- Lets construct a probability distribution for this class.

Mean and variance of a probability distribution

- The mean or **expected value** of a probability distribution is:

$$\mu = \sum x_i P(x_i)$$

- The variance of a probability distribution is given by:

$$\sigma^2 = \sum [(x_i - \mu)^2 P(x_i)]$$

$$\sigma^2 = \sum (x_i^2 P(x_i)) - \mu^2$$

- Try calculating the mean, variance, and standard deviation for the previous example.

Mean and variance of a probability distribution

- The mean or **expected value** of a probability distribution is:

$$\mu = \sum x_i P(x_i)$$

- The variance of a probability distribution is given by:

$$\sigma^2 = \sum [(x_i - \mu)^2 P(x_i)]$$

$$\sigma^2 = \sum (x_i^2 P(x_i)) - \mu^2$$

- Try calculating the mean, variance, and standard deviation for the previous example.

Mean and variance of a probability distribution

- The mean or **expected value** of a probability distribution is:

$$\mu = \sum x_i P(x_i)$$

- The variance of a probability distribution is given by:

$$\sigma^2 = \sum \left[(x_i - \mu)^2 P(x_i)\right]$$

$$\sigma^2 = \sum \left(x_i^2 P(x_i)\right) - \mu^2$$

- Try calculating the mean, variance, and standard deviation for the previous example.

Mean and variance of a probability distribution

- The mean or **expected value** of a probability distribution is:

$$\mu = \sum x_i P(x_i)$$

- The variance of a probability distribution is given by:

$$\sigma^2 = \sum \left[(x_i - \mu)^2 P(x_i) \right]$$

$$\sigma^2 = \sum \left(x_i^2 P(x_i) \right) - \mu^2$$

- Try calculating the mean, variance, and standard deviation for the previous example.

Mean and variance of a probability distribution

- The mean or **expected value** of a probability distribution is:

$$\mu = \sum x_i P(x_i)$$

- The variance of a probability distribution is given by:

$$\sigma^2 = \sum \left[(x_i - \mu)^2 P(x_i)\right]$$

$$\sigma^2 = \sum \left(x_i^2 P(x_i)\right) - \mu^2$$

- Try calculating the mean, variance, and standard deviation for the previous example.

Mean and variance of a probability distribution

- The mean or **expected value** of a probability distribution is:

$$\mu = \sum x_i P(x_i)$$

- The variance of a probability distribution is given by:

$$\sigma^2 = \sum \left[(x_i - \mu)^2 P(x_i) \right]$$

$$\sigma^2 = \sum \left(x_i^2 P(x_i) \right) - \mu^2$$

- Try calculating the mean, variance, and standard deviation for the previous example.

Covariance

- To measure the variance of combinations of RVs, need to know the covariance.
- **Covariance**: measure of how two RVs move together.
- Notation:
  - σ_{xy}: population covariance.
  - s_{xy}: sample covariance.
- Formulas:

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

- Interpretations:
  - When covariance is negative, variables move in opposite directions.
  - When covariance is positive, variables move in same direction.

Covariance

- To measure the variance of combinations of RVs, need to know the covariance.
- **Covariance**: measure of how two RVs move together.
- Notation:
  - σ_{xy}: population covariance.
  - s_{xy}: sample covariance.
- Formulas:

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

- Interpretations:
  - When covariance is negative, variables move in opposite directions.
  - When covariance is positive, variables move in same direction.

Covariance

- To measure the variance of combinations of RVs, need to know the covariance.
- **Covariance**: measure of how two RVs move together.
- Notation:
  - σ_{xy}: population covariance.
  - s_{xy}: sample covariance.
- Formulas:

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

- Interpretations:
  - When covariance is negative, variables move in opposite directions.
  - When covariance is positive, variables move in same direction.

Covariance

- To measure the variance of combinations of RVs, need to know the covariance.
- **Covariance**: measure of how two RVs move together.
- Notation:
  - σ_{xy}: population covariance.
  - s_{xy}: sample covariance.
- Formulas:

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

- Interpretations:
  - When covariance is negative, variables move in opposite directions.
  - When covariance is positive, variables move in same direction.

Covariance

- To measure the variance of combinations of RVs, need to know the covariance.
- **Covariance**: measure of how two RVs move together.
- Notation:
  - σ_{xy}: population covariance.
  - s_{xy}: sample covariance.
- Formulas:

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

- Interpretations:
  - When covariance is negative, variables move in opposite directions.
  - When covariance is positive, variables move in same direction.

Covariance

- To measure the variance of combinations of RVs, need to know the covariance.
- **Covariance**: measure of how two RVs move together.
- Notation:
  - σ_{xy}: population covariance.
  - s_{xy}: sample covariance.
- Formulas:

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

- Interpretations:
  - When covariance is negative, variables move in opposite directions.
  - When covariance is positive, variables move in same direction.

Covariance

- To measure the variance of combinations of RVs, need to know the covariance.
- **Covariance**: measure of how two RVs move together.
- Notation:
  - σ_{xy}: population covariance.
  - s_{xy}: sample covariance.
- Formulas:

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

- Interpretations:
  - When covariance is negative, variables move in opposite directions.
  - When covariance is positive, variables move in same direction.

Covariance

- To measure the variance of combinations of RVs, need to know the covariance.
- **Covariance**: measure of how two RVs move together.
- Notation:
  - σ_{xy}: population covariance.
  - s_{xy}: sample covariance.
- Formulas:

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

- Interpretations:
  - When covariance is negative, variables move in opposite directions.
  - When covariance is positive, variables move in same direction.

Covariance

- To measure the variance of combinations of RVs, need to know the covariance.
- **Covariance**: measure of how two RVs move together.
- Notation:
  - σ_{xy}: population covariance.
  - s_{xy}: sample covariance.
- Formulas:

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

- Interpretations:
  - When covariance is negative, variables move in opposite directions.
  - When covariance is positive, variables move in same direction.

Combining Random Variables

- Suppose two RVs are measured in the same units, can they be combined (added, subtracted, etc)?
- Who would want to do such a thing?
- Addition rule:
  - If $Z = X + Y$, then $E(Z) = E(X) + E(Y)$.
  - More generally, if $Z = aX + bY$, then $E(Z) = aE(X) + bE(Y)$.

Combining Random Variables

- Suppose two RVs are measured in the same units, can they be combined (added, subtracted, etc)?
- Who would want to do such a thing?
- Addition rule:
  - If $Z = X + Y$, then $E(Z) = E(X) + E(Y)$.
  - More generally, if $Z = aX + bY$, then $E(Z) = aE(X) + bE(Y)$.

Combining Random Variables

- Suppose two RVs are measured in the same units, can they be combined (added, subtracted, etc)?
- Who would want to do such a thing?
- Addition rule:
  - If $Z = X + Y$, then $E(Z) = E(X) + E(Y)$.
  - More generally, if $Z = aX + bY$, then $E(Z) = aE(X) + bE(Y)$.

Combining Random Variables

- Suppose two RVs are measured in the same units, can they be combined (added, subtracted, etc)?
- Who would want to do such a thing?
- Addition rule:
  - If $Z = X + Y$, then $E(Z) = E(X) + E(Y)$.
  - More generally, if $Z = aX + bY$, then $E(Z) = aE(X) + bE(Y)$.

Combining Random Variables

- Suppose two RVs are measured in the same units, can they be combined (added, subtracted, etc)?
- Who would want to do such a thing?
- Addition rule:
  - If $Z = X + Y$, then $E(Z) = E(X) + E(Y)$.
  - More generally, if $Z = aX + bY$, then $E(Z) = aE(X) + bE(Y)$.

Variance of Combinations

- Suppose $Z = X + Y$:

$$VAR(X+Y) \equiv \sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2\sigma_{xy}.$$

- Suppose $Z = aX + bY$:

$$VAR(aX+bY) \equiv \sigma_z^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}.$$

Variance of Combinations

- Suppose $Z = X + Y$:

$$VAR(X + Y) \equiv \sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2\sigma_{xy}.$$

- Suppose $Z = aX + bY$:

$$VAR(aX + bY) \equiv \sigma_z^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}.$$

Variance of Combinations

- Suppose $Z = X + Y$:

$$VAR(X + Y) \equiv \sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2\sigma_{xy}.$$

- Suppose $Z = aX + bY$:

$$VAR(aX + bY) \equiv \sigma_z^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}.$$

Variance of Combinations

- Suppose $Z = X + Y$:

$$VAR(X + Y) \equiv \sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2\sigma_{xy}.$$

- Suppose $Z = aX + bY$:

$$VAR(aX + bY) \equiv \sigma_z^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}.$$

Example

- Suppose you run a website business that produces and sells custom made furniture.
  - When someone submits an order online you must first build the furniture. The average time it takes to build the furniture is 9 business days and the standard deviation is 2.5 business days.
  - After you build the furniture, you must ship it. The average shipping time is 4 business days and the standard deviation is 3 business days.
  - What is the average elapsed time from the time the order is placed until the delivery?
  - What is the variance and standard deviation of the elapsed time from the time the order is placed until the delivery?

Example

- Suppose you run a website business that produces and sells custom made furniture.
  - When someone submits an order online you must first build the furniture. The average time it takes to build the furniture is 9 business days and the standard deviation is 2.5 business days.
  - After you build the furniture, you must ship it. The average shipping time is 4 business days and the standard deviation is 3 business days.
  - What is the average elapsed time from the time the order is placed until the delivery?
  - What is the variance and standard deviation of the elapsed time from the time the order is placed until the delivery?

Example

- Suppose you run a website business that produces and sells custom made furniture.
  - When someone submits an order online you must first build the furniture. The average time it takes to build the furniture is 9 business days and the standard deviation is 2.5 business days.
  - After you build the furniture, you must ship it. The average shipping time is 4 business days and the standard deviation is 3 business days.
  - What is the average elapsed time from the time the order is placed until the delivery?
  - What is the variance and standard deviation of the elapsed time from the time the order is placed until the delivery?

Example

- Suppose you run a website business that produces and sells custom made furniture.
  - When someone submits an order online you must first build the furniture. The average time it takes to build the furniture is 9 business days and the standard deviation is 2.5 business days.
  - After you build the furniture, you must ship it. The average shipping time is 4 business days and the standard deviation is 3 business days.
  - What is the average elapsed time from the time the order is placed until the delivery?
  - What is the variance and standard deviation of the elapsed time from the time the order is placed until the delivery?

Example

- Suppose you run a website business that produces and sells custom made furniture.
  - When someone submits an order online you must first build the furniture. The average time it takes to build the furniture is 9 business days and the standard deviation is 2.5 business days.
  - After you build the furniture, you must ship it. The average shipping time is 4 business days and the standard deviation is 3 business days.
  - What is the average elapsed time from the time the order is placed until the delivery?
  - What is the variance and standard deviation of the elapsed time from the time the order is placed until the delivery?