Random Variables and Probability Distributions

Mgmt 230: Introductory Statistics

1 Probability Distributions

1.1 Random Variables

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.

1.2 Probability Distribution

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.

1.3 Mean and variance

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.

2 Binomial distribution

2.1 Bernoulli Trials

Binomial distributions

- A **Bernoulli trial** results in a random variable that can only result in success ($x = 1$) or failure ($x = 0$).
  - Example: an outcome of heads for a single coin flip is a Bernoulli trial.
- A **binomial distribution** is the probability distribution for the number of successes in a fixed number of trials.
- Requirements for a binomial distribution:
  - The experiment must have a fixed number of Bernoulli trials.
  - The trials must be independent.
  - The probability of success must be the same for each trial.
- Example: what are the possible outcomes for the number of successes in 5 coin flips?

2.2 Binomial distribution

Binomial probability distribution

- The binomial probability distribution is given by,

$$P(x) = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}$$

  - n: number of trials.
  - $P(x)$: the probability of x number of successes.
  - p is the probability of success for a single Bernoulli trial.

- Calculate the probability distribution of the 5 coin flip experiment.
- Verify $\sum P(x_i) = 1$.
- Calculate the expected value for the number of heads.
- Calculate the variance and standard deviation.

Binomial Distribution

- Recall the mean of a probability distribution:

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

- For the binomial distribution, this gets more simple:

$$\mu = np$$

- Recall the variance of a probability distribution:

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

- It gets simpler:

$$\sigma^2 = np(1-p)$$

2.3 Normal Approximation

Normal Approximation to a Binomial

- Motivation: Questions like,
- Suppose a study concluded that Burger King fills out 10% of its orders inaccurately. If a particular franchise makes 30,000 orders a month, what is the probability it will make more than 3,100 errors in orders?
- Compute P(x=3100), P(x=3101), P(x=3102), P(x=3103), ... , P(x=30,000).
- That's a ridiculous amount of computations.
- Can you suppose that the number of errors is *normally distributed* with mean equal to $\mu = np$, and variance $\sigma^2 = np(1-p)$?

Normal Approximation to a Binomial

- Can you suppose that the number of errors is *normally distributed* with mean equal to $\mu = np$, and variance $\sigma^2 = np(1-p)$?
- Well... the number of errors is truly a *binomial distribution*, not a normal distribution.
- And... the number of errors is a *discrete* random variable, but the normal distribution is for continuous random variables.
- But... the normal distribution will be a good approximation if,

$$np > 5 \text{ and } n(1-p)$$

3 Combining Random Variables

3.1 Addition Rule

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 $E(Z) = aE(X) + bE(Y)$.

3.2 Covariance

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.

3.3 Variance of Combinations

Variance of Combinations

- Suppose $Z = X + Y$:

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

- Suppose $Z = aX + bY$:

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

3.4 Portfolio Risk

Portfolio Risk

- Don't put all your eggs in one basket.
- Would it make more sense to put your money in two investments that are negatively correlated (meaning they have a negative covariance)?
- Example, suppose the variance for the quarterly return is equal to 15% for Investment X and 10% for Investment Y, and the covariance is equal to -8%. Suppose you invested have your money in each investment.
- Would it make more sense to put your money in two investments that are positively correlated (meaning they have a positive covariance)?
- Example, suppose the variance for the quarterly return is equal to 18% for Investment X and 12% for Investment Y, and the covariance is equal to 6%. Suppose you invested have your money in each investment.

4

Homework

- Binomial distribution: (not required) Section 5.3, pages 196-197: 5.18 through 5.25.
- Normal distribution: Section 6-2, problems 6.1, 6.3, 6.5, 6.6, 6.11.
- Combining RVs: Section 5.2, pages 187-189: 5.7 through 5.11.