Assignment 1 due Friday

- **Assignment 2 will be out soon.**
- **15/4% of your grade for each Assignment**
- **Reminder – 10% clickers, 5% chapter quizzes, 15% Assignments, 20% Mid-term exams, 10% Labs, 10% Project and 30% Final Exam**

Ass –ump-TIONS

- I see a linear model
- I see assumptions in epsilon

X2

- Independent Identically Distributed
- Normal Zero sigma squared.

$$\varepsilon_i \overset{iid}{\sim} N(0,\sigma^2)$$

P-value

- Left side:
- **P-value**

Middle:

Getting so small

Right side:

Kick Ho

- Everyone:
- **OUT the door**

Exam 1

- Chapters 1-5
- When?
- SEE canvas

Have you given your data set info on CANVAS?

A) Yes

B) No

Counts toward clicker grade

Mutually exclusive events, A and B: $P(A \cap B) =?$

If A and B are mutually exclusive then

$$P(A \cup B) = P(A) + P(B)$$

A) True

B) False

Evaluate 4!

The point estimate for β_1 is?

A) 0.2134

B) 2.4264

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.2134     0.4390   0.486    0.632
RATIO         2.4264     0.2283  10.630 3.92e-10 ***
---
```

$\beta_1 \neq 0$

A) True
B) False

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.2134     0.4390   0.486    0.632
RATIO         2.4264     0.2283  10.630 3.92e-10 ***
---
```

Chapter 4

Discrete Random Variables

Independence

$P(A \cap B) = P(A|B)P(B) = P(A)P(B)$ if and only if A and B are independent

Expectation Theorems

$$E(X) = \sum_i x_i \, p(X = x_i), Defn$$

$$E(cX) = cE(X)$$

$$E(c) = c$$

$$E(g_1(X) + g_2(X) + \ldots + g_n(X)) = E(g_1(X)) + E(g_2(X)) + \ldots + E(g_n(X))$$

NB!

$$E[g(X)] = \sum_i g(x_i)p(X = x_i)$$

Know this definition!

$$\sigma^2 = E[(X-\mu)^2] = \sum_i (x_i - \mu)^2 p(X = x_i)$$

Bayes’ Rule (Wikipedia example)

Suppose a drug test is 99% sensitive and 99% specific. That is, the test will produce 99% true positive results for drug users and 99% true negative results for non-drug users. Suppose that 0.5% of people are users of the drug. **If a randomly selected individual tests positive, what is the probability he or she is a user**?

This is the calculation!

$$P(\text{User}|+) = \frac{P(+|\text{User})P(\text{User})}{P(+|\text{User})P(\text{User}) + P(+|\text{Non-user})P(\text{Non-user})}$$

$$= \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.01 \times 0.995}$$

$$\approx 33.2\%$$

0.99 +0.01=1

$$\boldsymbol{P(-|Non.user) + P(+|Non.user) = 1}$$

Bernoulli Trials

- Each trial results in one of two outcomes (mutually exclusive), S and F
- No other outcomes
- $p(S) = p \text{ and } p(F) = q, p + q = 1$
- Y=1 when there is a S and Y=0 when there is a F

Bernoulli Probability distribution

$$p(y) = p^y q^{1-y}, y = 0,1$$

Calculate mean and variance

$$E[Y] = \sum_{y} y p(y) = \mu$$

$$\sigma^2 = \sum_{y} (y - \mu)^2 p(y)$$

Bernouli

- See BBD

Binomial distribution

- n Bernoulli trials
- Two possible outcomes per trial S or F
- $p(S) = p \; and \; p(F) = q, p + q = 1$
- Trials are independent
- Y is the number of successes in n trials

Binomial Probability Distribution

$$p(y) = \binom{n}{y} p^y q^{n-y}, y = 0,1,2,\ldots,n$$

$$\mu = np$$

$$\sigma^2 = npq$$

Binomial example

- 10% of computers have viruses. If a sample of size 10 computers were inspected what is the probability x=2 would have a virus?

See BBD

Project

- Start looking for an interesting data set suitable for linear regression
  - Internet
  - Library books
  - Text book chapter 10

The Geometric distribution

$$p(y) = pq^{y-1}, (y = 1, 2, \ldots)$$

$$\mu = \frac{1}{p}, \sigma^2 = \frac{q}{p^2}$$

- p=probability of success on a single Bernoulli trial
- $p + q = 1$
- Y=Number of trials until the first success,

The Hyper-geometric distribution

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}, y = Max[0, n-(N-r)], ..., Min(r, n)$$

$$\mu = \frac{nr}{N}, \sigma^2 = \frac{r(N-r)n(N-n)}{N^2(N-1)}$$

- N=Total number of elements
- r=Number of S's in the N elements
- n=Number of elements drawn
- Y=Number of S's drawn in the n elements

Characteristics of a hyper-geometric distribution

- The experiment consists of randomly drawing n elements without replacement from a set of N elements, r of which are S's and (N-r) of which are failures
- The sample size n is large relative to the number of elements in the population i.e. $n/N > 0.05$
- The hyper-geometric random variable Y is the number of S's in the draw of n elements.

Binomial example

- 10% of computers have viruses. If a sample of size 10 computers were inspected what is the probability x=2 would have a virus?

See BBD

The Multinomial Distribution

$$p(y_1, y_2, ..., y_k) = \frac{n!}{y_1! y_2! .... y_k!} p_1^{y_1} p_2^{y_2} ... p_k^{y_k}$$

$p_i =$ probability of outcome i on a single trial

$p_1 + p_2 + ... + p_k = 1$

$n = y_1 + y_2 + ... + y_k =$ Number of trials

$y_i =$ Number of occurrences of outcome i in n trials

$\mu_i = np_i$ and $\sigma_i^2 = np_i(1 - p_i)$

- The experiment consists of n identical trials
- There are k possible outcomes
- The probabilities remain constant from trial to trial
- Trials are independent
- Y's are of interest (one per category)

Example of the multinomial

- N=103 sparkplug are found defective. There are 5 production lines.

A	B	C	D	E
15	27	31	19	11

What is the probability that there are 2,3,4,3,5 sparkplug failures

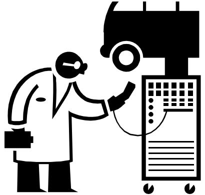
The Negative Binomial

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r}, (y = r, r+1, r+2, \ldots)$$

$$\mu = \frac{r}{p}, \sigma^2 = \frac{rq}{p^2}$$

- p=probability of success on a single Bernoulli trial
- $p + q = 1$
- Y=Number of trials until the r^{th} success,

Example 4.13 page 143



The Geometric distribution

$$p(y) = pq^{y-1}, (y = 1, 2, \ldots)$$

$$\mu = \frac{1}{p}, \sigma^2 = \frac{q}{p^2}$$

- p=probability of success on a single Bernoulli trial
- $p + q = 1$
- Y=Number of trials until the first success,

The Hyper-geometric distribution

$$p(y)=\frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}, y=Max[0,n-(N-r)],...,Min(r,n)$$

$$\mu=\frac{nr}{N}, \sigma^2=\frac{r(N-r)n(N-n)}{N^2(N-1)}$$

- N=Total number of elements
- r=Number of S's in the N elements
- n=Number of elements drawn
- Y=Number of S's drawn in the n elements

Example 4.15 page 148

The Poisson

$$p(y) = \frac{\lambda^y e^{-\lambda}}{y!}, y = 0,1,2,\ldots$$

$$\mu = \lambda$$

$$\sigma^2 = \lambda$$

Example 4.18

- Mean=2.5 cracks/specimen of concrete

$$k^{th}\ moment\ about\ zero$$

$$\mu_k' = E\left(Y^k\right), k = 1,2,3\ \ldots$$

Moments (Definitions)

$$\mu_k^{'} = E(Y^k), (\mathrm{k} = 1,2,...) \text{ This is the kth moment about zero}$$

$$\mu_k = E[(Y-\mu)^k], \text{ this is the kth moment about } \mu$$

Moment Generating Function

$$m_X(t) = E(e^{Xt})$$

Moment generating theorem

$$\mu_k' = \left.\frac{d^k m(t)}{dt^k}\right]_{t=0}$$

Prove for a Bernoulli

$$m(t) = \left(pe^{t} + q\right)$$

Prove for a Binomial

$$m(t) = \left(pe^t + q\right)^n$$

Assignment 2

- **Coming soon**
- **5% of your grade for each Assignment**
- **Reminder – 10% quizzes, 5% chapter quizzes, 15% Assignments, 20% Mid-term exams, 10% Labs, 10% Project and 30% Final Exam**

Exam 1 chapters 1-5

- Be ready for the exam
- Coming soon
- See CANVAS for considerable hints!!

Chapter 4

Discrete Random Variables

How many discrete distributions do we study?

- A) 6
- B) 7
- C) 8
- D) 10

- Task 5
  - Use R to calculate
    - $\binom{8}{4}$ – hint: Try choose()
    - $P(Y > 4), Y \sim Pois(\lambda = 2)$
  - Some more calculations in R
    - $P(Y = 10), Y \sim NegBin(p = 0.4, r = 3)$
    - $P(Y \leq 8), Y \sim Bin(n = 15, p = 0.4)$

Bernoulli Trials

- Each trial results in one of two outcomes (mutually exclusive), S and F
- No other outcomes
- $p(S) = p \ and \ p(F) = q, p + q = 1$
- Y=1 when there is a S and Y=0 when there is a F

Bernoulli Probability distribution

$$p(y) = p^{y} q^{1-y}, y = 0,1$$

Calculate mean and variance

$$E[Y] = \sum_{y} y p(y) = \mu$$

$$\sigma^2 = \sum_{y} (y - \mu)^2 p(y)$$

Binomial distribution

- n Bernoulli trials
- Two possible outcomes per trial S or F
- $p(S) = p \text{ and } p(F) = q, p + q = 1$
- Trials are independent
- Y is the number of successes in n trials

Binomial Probability Distribution

$$p(y) = \binom{n}{y} p^y q^{n-y}, y = 0, 1, 2, \ldots, n$$

$$\mu = np$$

$$\sigma^2 = npq$$

The Multinomial Distribution

$$p(y_1, y_2, ..., y_k) = \frac{n!}{y_1! y_2! .... y_k!} p_1^{y_1} p_2^{y_2} ... p_k^{y_k}$$

$p_i =$ probability of outcome i on a single trial

$p_1 + p_2 + ... + p_k = 1$

$n = y_1 + y_2 + ... + y_k =$ Number of trials

$y_i =$ Number of occurrences of outcome i in n trials

$\mu_i = np_i$ and $\sigma_i^2 = np_i(1 - p_i)$

- The experiment consists of n identical trials
- There are k possible outcomes
- The probabilities remain constant from trial to trial
- Trials are independent
- Y's are of interest

The Negative Binomial

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r}, (y = r, r+1, r+2, \ldots)$$

$$\mu = \frac{r}{p}, \sigma^2 = \frac{rq}{p^2}$$

- p=probability of success on a single Bernoulli trial
- $p + q = 1$
- Y=Number of trials until the r^{th} success,

The Geometric distribution

$$p(y) = pq^{y-1}, (y = 1,2,\ldots)$$

$$\mu = \frac{1}{p}, \sigma^2 = \frac{q}{p^2}$$

- p=probability of success on a single Bernoulli trial
- $p + q = 1$
- Y=Number of trials until the first success,

The Hyper-geometric distribution

$$p(y)=\frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}, y=Max[0,n-(N-r)],...,Min(r,n)$$

$$\mu=\frac{nr}{N}, \sigma^2=\frac{r(N-r)n(N-n)}{N^2(N-1)}$$

- N=Total number of elements
- r=Number of S's in the N elements
- n=Number of elements drawn
- Y=Number of S's drawn in the n elements

Characteristics of a hyper-geometric distribution

- The experiment consists of randomly drawing n elements without replacement from a set of N elements, r of which are S's and (N-r) of which are failures
- The sample size n is large relative to the number of elements in the population i.e. $n/N > 0.05$
- The hyper-geometric random variable Y is the number of S's in the draw of n elements.

The Poisson

$$p(y) = \frac{\lambda^y e^{-\lambda}}{y!}, y = 0,1,2,\ldots$$

$$\mu = \lambda$$

$$\sigma^2 = \lambda$$

Characteristics of Poisson

- Experiment consists of counting the number of times Y a particular event occurs over a unit time, area or volume … (unit of measure)
- The probability that an event occurs in a given **unit** time … is the same for all **units**
- The number of events that occur in one unit of time … is independent of the number that occur in other units

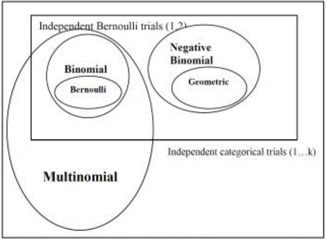
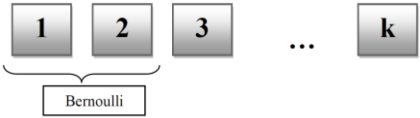


Figure 1. Venn diagram showing commonalities and differences in some of the distributions

The multinomial is related to the binomial in that each trial can result in one of k categories rather than two in the case of the binomial, the trials are not repeated independent Bernoulli but repeated independent categorical distributions, the Bernoulli then is seen as a special case of the categorical distribution. This distribution along with all the Bernoulli based distributions utilize at least one probability of success (k of them in a Multinomial).



Are there categorical trials?

a. Yes
    i. Is the number of trials fixed?
        1. Yes
            a. Are there more than two categories?
                i. Yes – Multinomial
                ii. No
                    1. Are there 2 or more trials
                        a. Yes – Binomial
                        b. No -- Bernoulli
        2. No
            a. Trials till first success?
                i. Yes – Geometric
                ii. No -- Negative Binomial
b. No
    i. Is there a constant rate?
        1. Yes – Poisson
        2. No – Hyper-geometric

$$p_i \ldots or \ldots p$$

Both used with categorical or Bernoulli trials (NOT Hyper-geometric or Poisson)

$$P(A \cup A^C) = ?$$

- A) 0
- B) 1
- C) ½

$$P(A|B) = \frac{P(A \cap B)}{P(A)}$$

- A) TRUE
- B) FALSE

Find $P(X = 3)$

A card is chosen from a shuffled 52 card deck (4 suites), replaced and another is drawn, etc

Let X= number of Spades, after 10 trials?

What is this experiment best described as?

A) Bernoulli B) Binomial C) Multinomial D) Hypergeometric

Find $P(X = 3)$

A card is chosen from a shuffled 52 card deck (4 suites), NOT replaced and another is drawn, etc

Let X= number of Spades, after 10 trials?

What is this experiment best described as?

A) Bernoulli B) Binomial C) Multinomial D) Hypergeometric

Find $P(X \leq 2)$

3. A batch of 20 integrated circuit chips contains 20% defective chips. A sample of 10 is drawn at random.

$X =$ the number of defective chips in the sample.

A) Binomial B) Multinomial C) Hypergeometric D) Neg Binomial

Best Answer!!

A wallet contains 3 \$100 bills and 5 \$1 bills. You randomly choose 4 bills. What is the probability that you will choose exactly 2 \$100 bills?

This is a Binomial!

A) True

B) False

Best Answer!

Suppose a biased coin comes up heads with probability 0.3 when tossed. What is the probability of 3 heads in 5 independent trials?

A) Binomial

B) Hyper-Geometric

Continuous Random Variables

Chapter 5

Chapter 5

CONTENTS

How to distinguish between Discrete and Continuous rvs

Many random variables observed in real life are not discrete random variables because the number of values that they can assume is not countable. For example, the waiting time Y (in minutes) at a traffic light could, in theory, assume any of the uncountably infinite number of values in the interval $0 < Y < \infty$. The daily rainfall at some location, the strength (in pounds per square inch) of a steel bar, and the intensity of sunlight at a particular time of the day are other examples of random variables that can assume any one of the uncountably infinite number of points in one or more intervals on the real line. In contrast to discrete random variables, such variables are called **continuous random variables**.

The preceding discussion identifies the difference between discrete and continuous random variables, but it fails to point to a practical problem. It is impossible to assign a finite amount of probability to each of the uncountable number of points in a line interval in such a way that the sum of the probabilities is 1. Therefore, the distinction between discrete and continuous random variables is usually based on the difference in their **cumulative distribution functions**.

Definition 5.1

The **cumulative distribution function $F(y_0)$** for a random variable Y is equal to the probability

$$F(y_0) = P(Y \leq y_0), -\infty < y_0 < \infty$$

For a discrete random variable, the cumulative distribution function is the cumulative sum of $p(y)$, from the smallest value that Y can assume, to a value of y_0. For example, from the cumulative sums in Table 2 of Appendix B, we obtain the following values of $F(y)$ for a binomial random variable with $n = 5$ and $p = .5$:

$$F(0) = P(Y \leq 0) = \sum_{y=0}^{0} p(y) = p(0) = .031$$

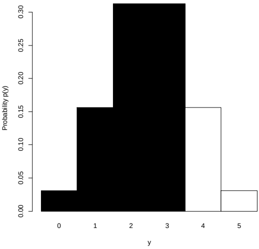
$$F(1) = P(Y \leq 1) = \sum_{y=0}^{1} p(y) = .188$$

$$F(2) = P(Y \leq 2) = \sum_{y=0}^{2} p(y) = .500$$

$$F(3) = P(Y \leq 3) = .812$$

$$F(4) = P(Y \leq 4) = .969$$

$$F(5) = P(Y \leq 5) = 1$$


Probability p(y)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
1
2
3
4
5
y

A graph of $p(y)$ is shown in Figure 5.1. The value of $F(y_0)$ is equal to the sum of the areas of the probability rectangles from $Y = 0$ to $Y = y_0$. The probability $F(3)$ is shaded in the figure.

A graph of the cumulative distribution function for the binomial random variable with $n = 5$ and $p = .5$, shown in Figure 5.2, illustrates an important property of the cumulative distribution functions for all discrete random variables: *They are step functions*. For example, $F(y)$ is equal to .031 until, as Y increases, it reaches $Y = 1$. Then $F(y)$ jumps abruptly to $F(1) = .188$. The value of $F(Y)$ then remains constant as Y increases until Y reaches $Y = 2$. Then $F(y)$ rises abruptly to $F(2) = .500$. Thus, $F(y)$ is a discontinuous function that jumps upward at a countable number of points ($Y = 0, 1, 2, 3$, and 4).

In contrast to the cumulative distribution function for a discrete random variable, the cumulative distribution function $F(y)$ for a continuous random variable is a

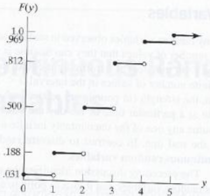


FIGURE 5.2
Cumulative distribution function $F(y)$ for a binomial random variable ($n = 5$, $p = 5$)

Definition 5.2

A continuous random variable Y is one that has the following three properties:

1. Y takes on an uncountably infinite number of values in the interval $(-\infty, \infty)$.
2. The cumulative distribution function, $F(y)$, is continuous.
3. The probability that Y equals any one particular value is 0.

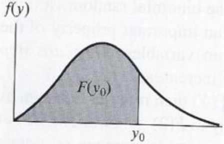


FIGURE 5.4
Density function $f(y)$ for a continuous random variable

$$f(y) = \frac{dF(y)}{dy}$$

$$F(y) = \int_{-\infty}^{y} f(t)\, dt$$

Properties of a density function

1.) $f(y) \geq 0$

2.) $\int_{-\infty}^{\infty} f(y)dy = F(\infty) = 1$

3.) $P(a < Y < b) = \int_{a}^{b} f(y)dy = F(b) - F(a)$

Expected values (Definitions)

$$E(Y) = \int_{-\infty}^{\infty} y f(y) dy$$

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

Expected value theorems

$$E(c) = c$$
$$E(cY) = cE(Y)$$
$$E[g_1(Y) + g_2(Y) + \ldots + g_k(Y)] = E(g_1(Y)) + \ldots + E(g_k(Y))$$

Y continuous and $E(Y) = \mu$

$$\sigma^2 = V(Y) = E[(Y-\mu)^2] = E(Y^2) - \mu^2$$

Integration formulae

$$\int y^m e^{ay} dy = \frac{y^m e^{ay}}{a} - \frac{m}{a} \int y^{m-1} e^{ay} dy$$

Theorem (V=variance)

$$V(cY) = c^2 V(Y)$$

$$V(c+Y) = V(Y)$$

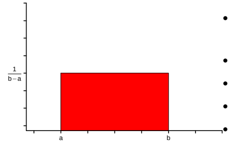
The uniform probability distribution

$$f(y) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le y \le b \\ 0 & \text{elsewhere} \end{cases}$$

$$\mu = \frac{a+b}{2} \quad \sigma^2 = \frac{(b-a)^2}{12}$$

Plotting the uniform

Uniform distribution



Code

- plot(
  1:10,type="n",axes=FALSE,ylab="",xlab="")
- polygon(c(2,2,6,6),c(0,4,4,0),col="Red")
- axis(1,0:8,c("","","a","","","","b","",""))
- axis(2,0:8,c("","","","",
- expression(frac(1,(b-a)),"","","",""))

Project

- Start looking for an interesting data set suitable for linear regression
  - Internet
  - Library books
  - Text book chapter 10