CS544, Fundamentals of Analysis Homework 4

Jefferson Parker

[japarker@bu.edu](mailto:japarker@bu.edu)

**Part 1, Binomial Distribution**

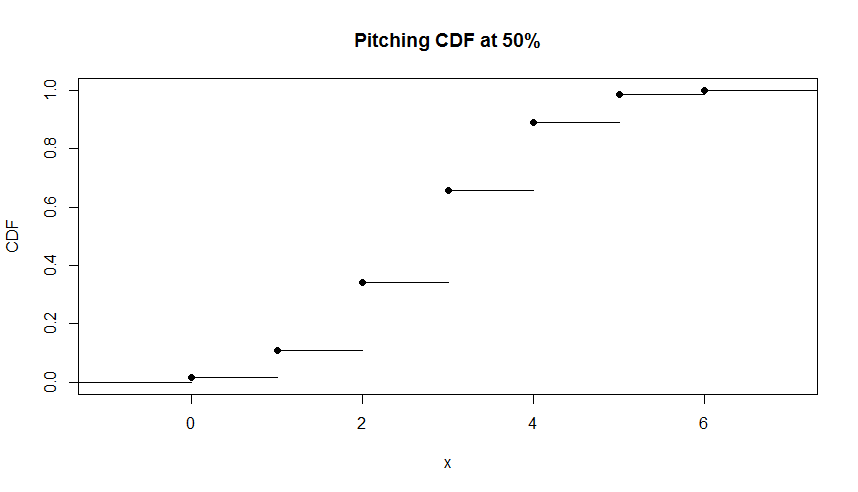
For a pitcher in baseball having a 50% chance of getting a strike out, using the binomial distribution:

1. Compute and plot the probability distribution for striking out the next 6 batters.

The probability of striking out the next six batters is 1.6% (0.015625).

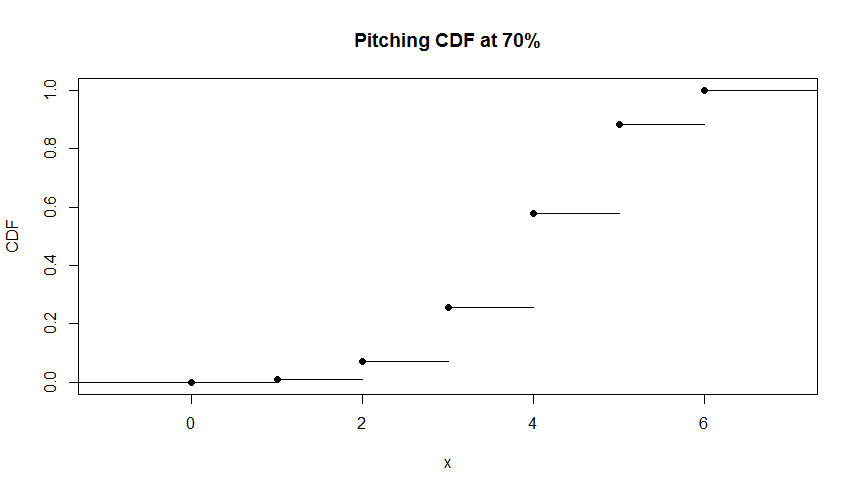
1. Plot the CDF of a).

The CDF plot is:



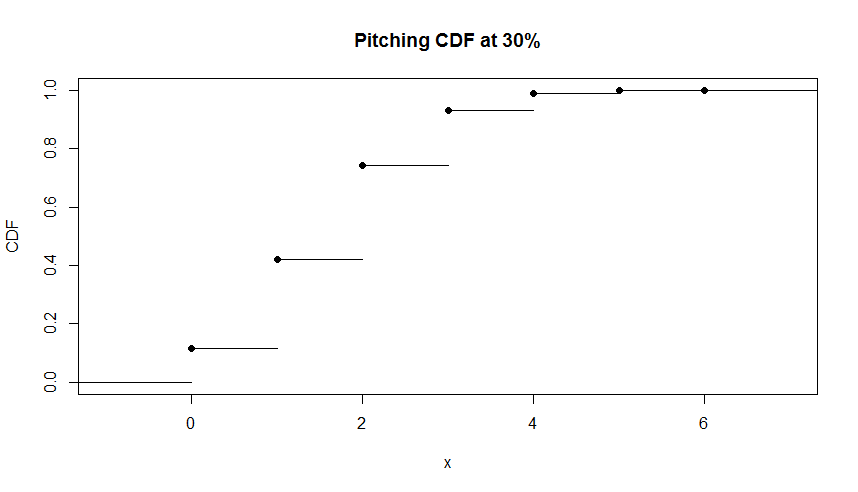
1. Repeat a) and b) assuming a 70% chance of getting a strike out.

The probability of striking out the next six batters is 11.8% (0.117649). The CDF plot is:



1. Repeat a) and b) assuming a 30% chance of getting a strike out.

The probability of striking out the next six batters is 0.07% (0.000729). The CDF plot is:



1. Make inferences from the shapes of the distributions.

From the curves of the strikeout CDF’s, we can see that as the probability of achieving a strikeout increases from 30% to 70% it becomes increasingly difficult to strike out batters later in the lineup. With a 30% chance of achieving a strike out, it is more likely to get a strike out against batters 1 -4, but exceptionally unlikely to strike out batters 5-6. With a 50% chance of strikeout it is unlikely to strike out the first and last batters, with good chance of striking out the middle of the lineup. For a 70% chance of getting a strikeout, the first few batters are less likely to strike out, but the latter four batters are likely to be struck out.

**Part 2, Binomial Distribution**

For an airport where 80% of flights arrive on time, use the binomial distribution to determine:

1. What is the probability that four of the next ten flights will arrive on time?

The probability is 89.3%.

sum(dbinom(3:9, size = n, prob = p));

[1] 0.8925479

1. What is the probability that four or fewer of the next ten flights will arrive on time?

The probability is 0.6%.

pbinom(4, size = n, prob = p);

[1] 0.006369382

1. Compute the probability distribution for the next ten flights arriving on time.

The probability distribution for the next ten flights to arrive on time is (n = 10, p = 0.8):

dbinom(0:n, size = n, prob = p);

0.0000001024

0.0000040960

0.0000737280

0.0007864320

0.0055050240

0.0264241152

0.0880803840

0.2013265920

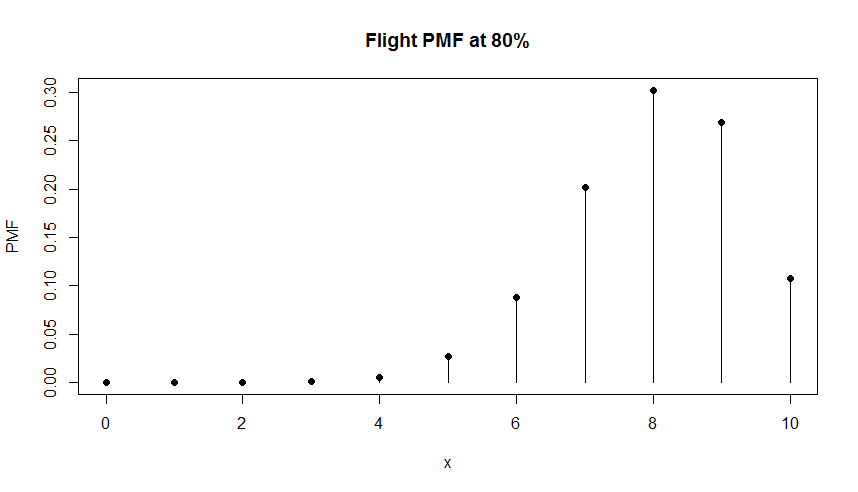
0.3019898880

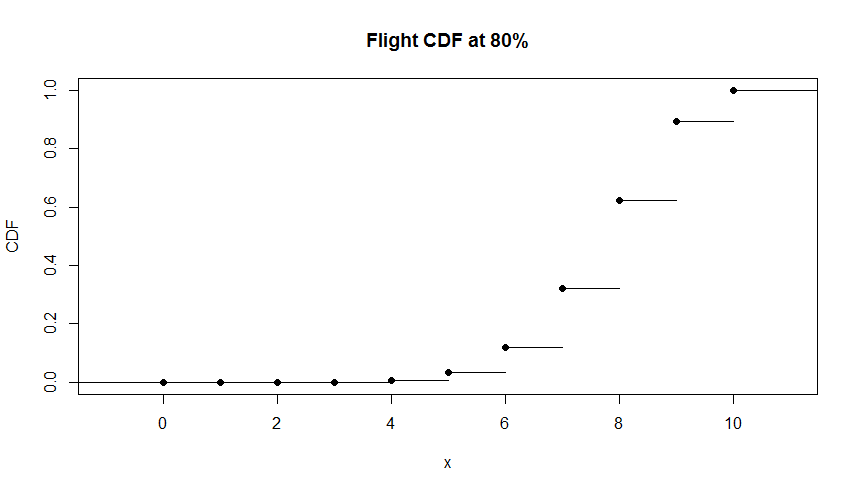
0.2684354560

0.1073741824

1. Show the PMF and CDF for the next ten flights.

The plots of PMF and CDF, respectively, for the flight arrivals are:





**Part 3, Poisson Distribution**

Using the Poisson distribution, calculate the following for a drive up window with an average of 10 cars arriving between 3:00 PM – 4:00 PM. For a lambda = 10:

1. What is the probability of serving exactly 3 cars?

The probability of serving exactly three cars is 0.7%

dpois(3, lambda = l);

[1] 0.007566655

1. What is the probability of serving at least 3 cars?

The probability of serving at least three cars (or more) is 99.2%.

1 - dpois(3, lambda = l);

[1] 0.9924333

1. What is the probability of serving 2 – 5 cars (inclusive)?

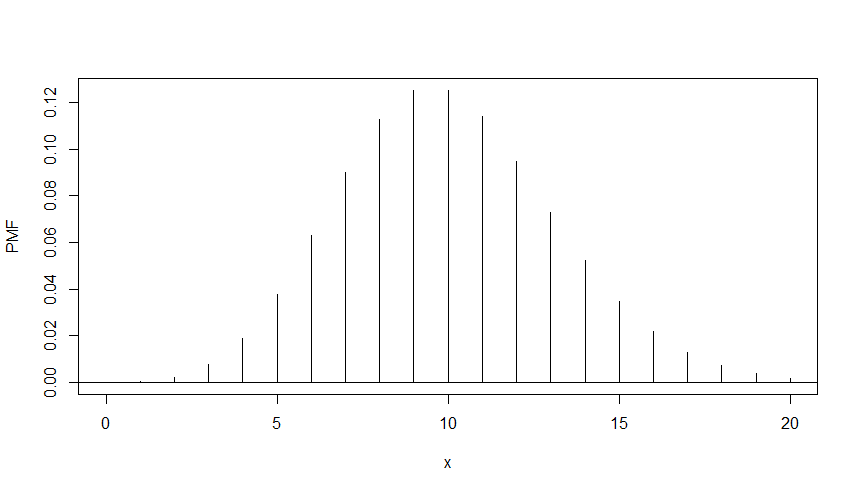
The probability of serving 2 to 5 cars is 6.4%.

ppois(5, lambda = l) - ppois(2, lambda = l);

[1] 0.06431657

1. Calculate and plot the PMF for the first 20 cars.

The plot of the PMF for the first 20 cars is:



The PMF for the first 20 cars is:

pmf <- dpois(0:20, lambda = l);

4.539993e-05

4.539993e-04

2.269996e-03

7.566655e-03

1.891664e-02

3.783327e-02

6.305546e-02

9.007923e-02

1.125990e-01

1.251100e-01

1.251100e-01

1.137364e-01

9.478033e-02

7.290795e-02

5.207710e-02

3.471807e-02

2.169879e-02

1.276400e-02

7.091109e-03

3.732163e-03

1.866081e-03

**Part 4, Uniform Distribution**

For a uniform distribution of exams between 60 – 100, inclusive:

1. What is the probability of scoring a 60? 80? 100?

The probability of scoring any of 60, 80 or 100 all have the same value of 2.5%.

dunif(60, min = 60, max = 100);

[1] 0.025

dunif(80, min = 60, max = 100);

[1] 0.025

dunif(100, min = 60, max = 100);

[1] 0.025

1. What is the mean and standard deviation of this distribution?

The mean is 80 and the standard deviation is 13.3.

(60 + 100)/2;

[1] 80

(sqrt((60 + 100)^2)/12);

[1] 13.33333

1. What is the probability of scoring at most a 70?

The probability of scoring up to 70 is 25%.

punif(70, min = 60, max = 100);

[1] 0.25

1. What is the probability of scoring greater than 80 (use lower.tail option)?

The probability of scoring over 80 is 50%.

punif(80, min = 60, max = 100, lower.tail = FALSE);

[1] 0.5

1. What is the probability of scoring between 90 – 100, inclusive?

The probability of scoring between 90 and 100 is 25%.

punif(100, min = 60, max = 100) - punif(90, min = 60, max = 100);

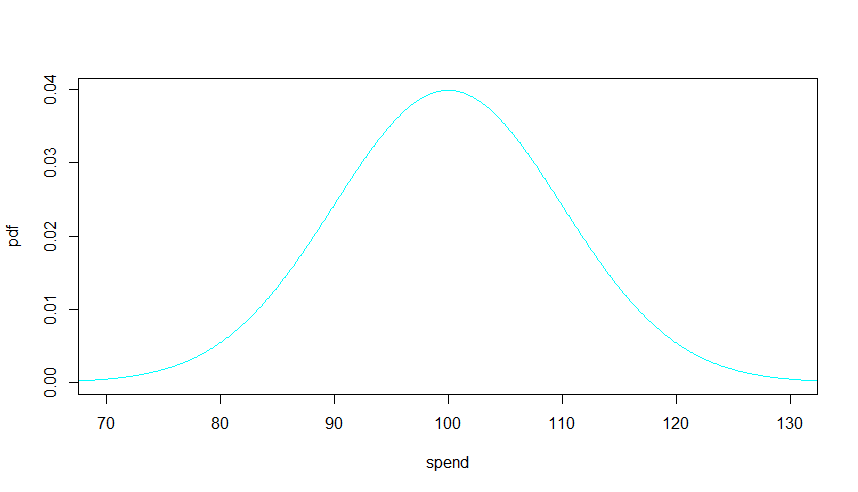
[1] 0.25

**Part 5, Normal Distribution**

For visitors to a theme park, where the average amount spent is $100 with a standard deviation of $10 and a normal distribution:

1. The PDF plot covering three standard deviations around the mean is:

The PDF plot is:



1. What is the probability of a randomly selected visitor spending more than $120?

The probability of a visitor spending at least $120 is 2.3%.

1 - pnorm(120, mean = mu, sd = sig);

[1] 0.02275013

1. What is the probability that a randomly selected visitor will spend between $80-$90, inclusive?

The probability that a visitor will spend $80-$90 is 13.6%

pnorm(90, mean = mu, sd = sig) - pnorm(80, mean = mu, sd = sig);

[1] 0.1359051

1. What are the probabilities of spending within one, two or three standard deviations respectively (assumed to be around the mean)?

The probability of spending within 1 SD is: 68.3%

2 SD: 95.4%

3 SD: 99.7%

pnorm(mu + 1\*sig, mean = mu, sd = sig) - pnorm(mu - 1\*sig, mean = mu, sd = sig);

[1] 0.6826895

> pnorm(mu + 2\*sig, mean = mu, sd = sig) - pnorm(mu - 2\*sig, mean = mu, sd = sig);

[1] 0.9544997

> pnorm(mu + 3\*sig, mean = mu, sd = sig) - pnorm(mu - 3\*sig, mean = mu, sd = sig);

[1] 0.9973002

1. Between what two values will the middle 90% of visitor spend fall?

90% of the spending per visitor will be between $87.18 and $112.82.

upper90 <- qnorm(0.9, mean = mu, sd = sig);

upper90;

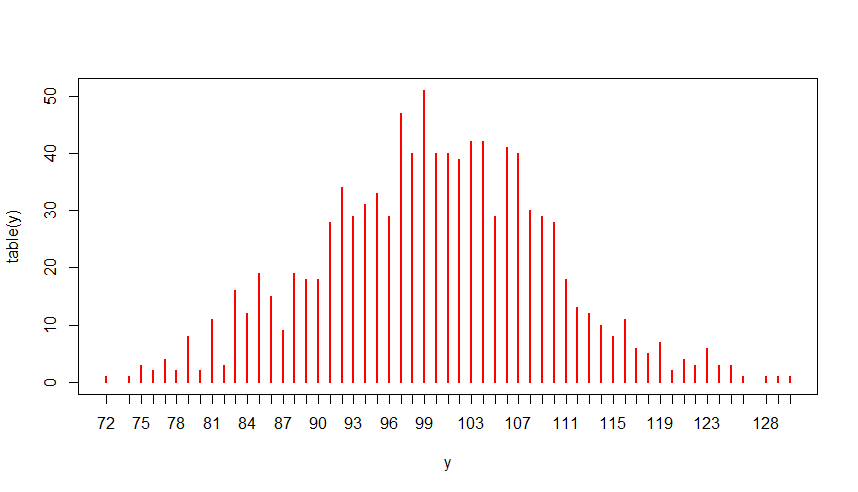
[1] 112.8155

mu - (upper90 - mu);

[1] 87.18448

1. Show a plot for 10,000 visitors using this distribution.

The plot for 10,000 visitors of this distribution will change with every instance. One plot for this data is:



**Part 6, Exponential Distribution**

Suppose a customer service provider receives 18 calls per hour.

1. What is the probability that the next call will arrive within 2 min?

The probability of the next call arriving within 2 min is 45.1%.

pexp(2/60, rate = 18);

[1] 0.4511884

1. What is the probability that the next call will arrive within 5 min?

The probability of the next call arriving within 5 min is 77.7%.

pexp(5/60, rate = 18);

[1] 0.7768698

1. What is the probability that the next call will arrive within 2 - 5 min?

The probability of the next call arriving within 2 to 5 min is 32.6%.

pexp(5/60, rate = 18) - pexp(2/60, rate = 18);

[1] 0.3256815

1. Plot the CDF of this distribution.

The CDF of this plot is

