Business Statistic

Week 5

Discrete Probability

Binomial and Poisson Distribution

Agenda

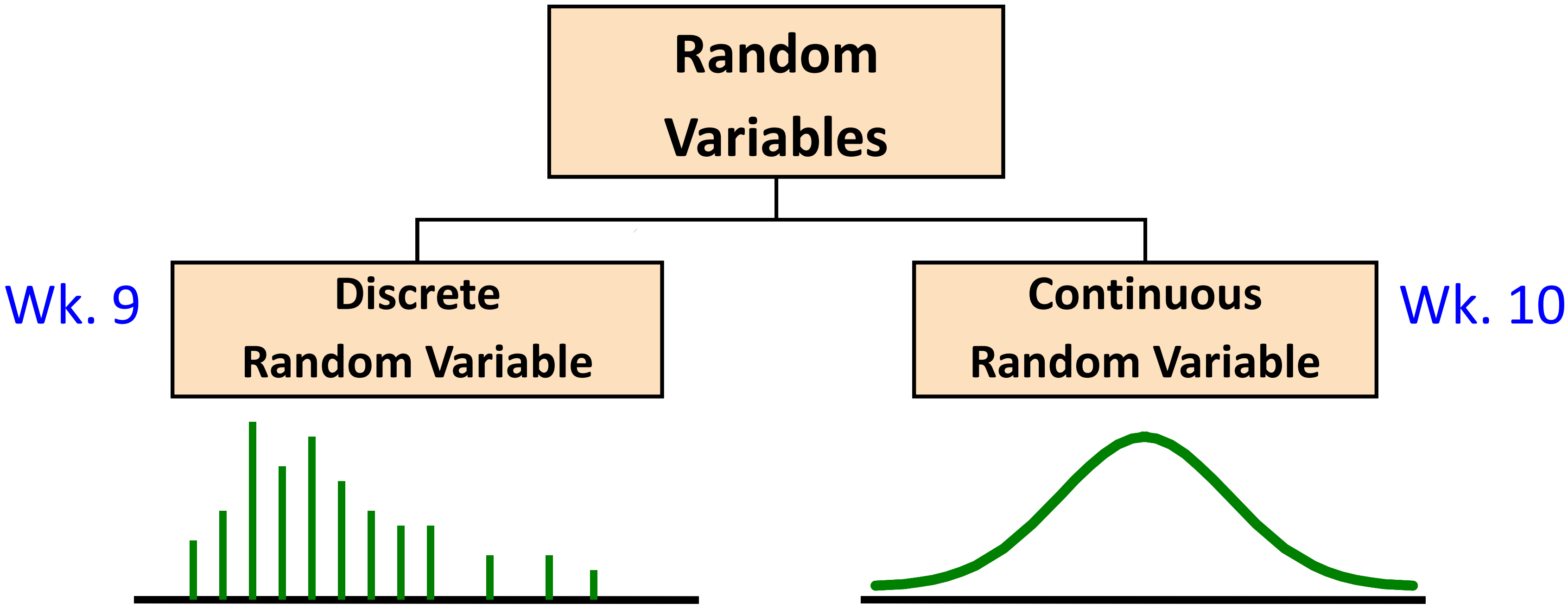
Time	Activity
50 minutes	Binomial Probability
50 minutes	Exercise
50 minutes	Poisson Probability
50 minutes	Exercise

Learning Objectives

In this chapter, you learn:

- To understand when to use Binomial and Poisson distributions
- To compute probabilities from the Binomial and Poisson distributions

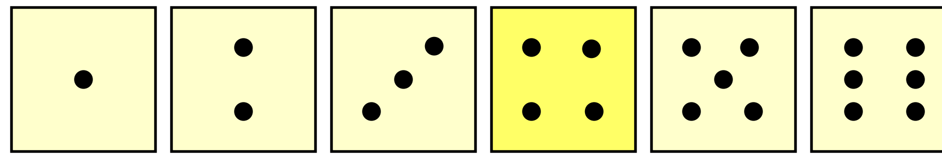
Random Variables



Discrete Random Variables

Can only assume a countable number of values

Examples:



- **Roll a die twice**

  Let X be the number of times 4 occurs

  (then X could be 0, 1, or 2 times)

- **Toss a coin 5 times.**

  Let X be the number of heads

  (then X = 0, 1, 2, 3, 4, or 5)

Probability Distribution For A Discrete Random Variable

A probability distribution for a discrete random variable is a mutually exclusive listing of all possible numerical outcomes for that variable and a probability of occurrence associated with each outcome.

Number of Classes Taken	Probability
2	0.20
3	0.40
4	0.24
5	0.16

Definitions
Random Variables

- A **random variable** represents a possible numerical value from an uncertain event.

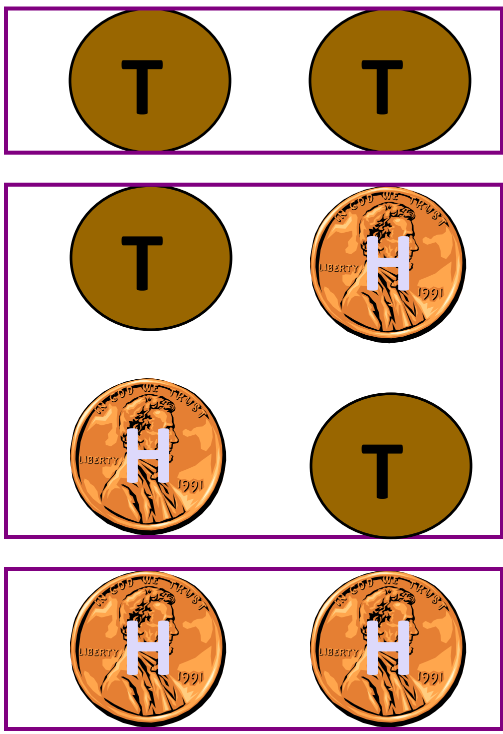
- **Discrete** random variables produce outcomes that come from a counting process (e.g. number of classes you are taking).

- **Continuous** random variables produce outcomes that come from a measurement (e.g. your annual salary, or your weight).

Example of a Discrete Random Variable Probability Distribution

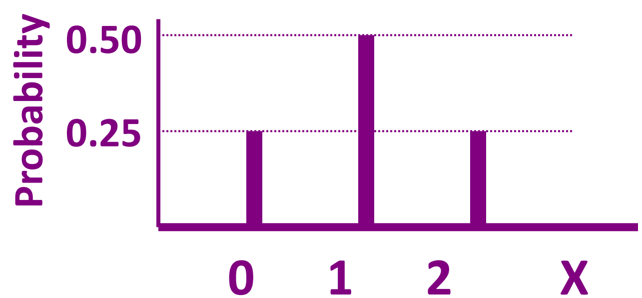
Experiment: Toss 2 Coins. Let X = # heads.

4 possible outcomes

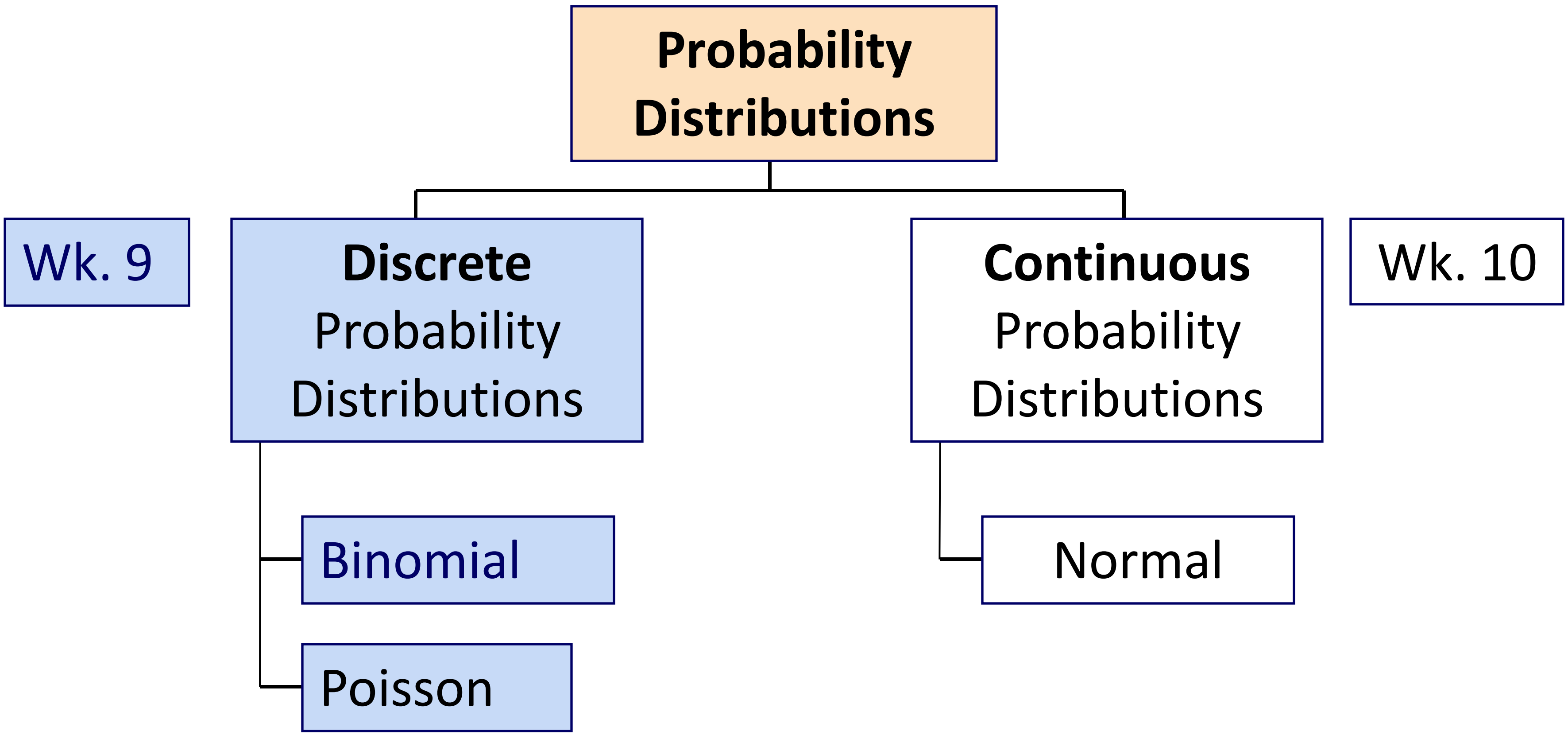


Probability Distribution

X Value	Probability
0	1/4 = 0.25
1	2/4 = 0.50
2	1/4 = 0.25



Probability Distributions



BINOMIAL PROBABILITY DISTRIBUTION

Binomial Probability Distribution

- A fixed number of observations, n
  - e.g., 15 tosses of a coin; ten light bulbs taken from a warehouse
- Each observation is categorized as to whether or not the "event of interest" occurred
  - e.g., head or tail in each toss of a coin; defective or not defective light bulb
  - Since these two categories are mutually exclusive and collectively exhaustive
    - When the probability of the event of interest is represented as π, then the probability of the event of interest not occurring is $1 - \pi$
- Constant probability for the event of interest occurring (π) for each observation
  - Probability of getting a tail is the same each time we toss the coin

Binomial Probability Distribution

(continued)

- Observations are independent
  - The outcome of one observation does not affect the outcome of the other
  - Two sampling methods deliver independence
    - Infinite population without replacement
    - Finite population with replacement

Possible Applications for the Binomial Distribution

- A manufacturing plant labels items as either defective or acceptable
- A firm bidding for contracts will either get a contract or not
- A marketing research firm receives survey responses of “yes I will buy” or “no I will not”
- New job applicants either accept the offer or reject it

Binomial Distribution Formula

$$P(X) = \frac{n\,!}{X\,!\,(n - X)!}\,\pi^{X}\,(1-\pi)^{n-X}$$

P(X) = probability of **X** events of interest in **n** trials, with the probability of an "event of interest" being **π** for each trial

X = number of "events of interest" in sample, (X = 0, 1, 2, ..., n)

n = sample size (number of trials or observations)

π = probability of "event of interest"

Example: Flip a coin four times, let x = # heads:

n = 4

π = 0.5

1 - π = (1 - 0.5) = 0.5

X = 0, 1, 2, 3, 4

Example

Suppose that you and your two friends go to Popeye’s, which last month filled approximately 85% of orders correctly. What is the probability that

a. all three orders will be filled correctly?
b. none of the three will be filled correctly?
c. at least two of the three will be filled correctly?

Example (Answer)

$n = 3$

$\pi = 0.85$

a. $P(X=3) = 0.6141$

b. $P(X=0) = 0.0034$

c. $P(X \geq 2) = P(X=2) + P(X=3)$

$$= 0.3251 + 0.6141 = 0.9392$$

Binomial Distribution Characteristics

- Mean

$$\mu = E(x) = n\pi$$

- Variance and Standard Deviation

$$\sigma^2 = n\pi(1 - \pi)$$

$$\sigma = \sqrt{n\pi(1 - \pi)}$$

Where n = sample size

π = probability of the event of interest for any trial

(1 – π) = probability of no event of interest for any trial

Example

For the problem in the example before, what are the mean and standard deviation of the binomial distribution used in (a) through (c)? Interpret these values.

Example (Answer)

- $\mu = n\pi = 3 * 0.85 =$ **2.55**
- $\sigma = \sqrt{n\pi(1-\pi)}$

  $= \sqrt{3*0.85(1-0.85)}$

  $= \sqrt{0.3825} =$ **0.6185**

The Binomial Distribution Using Binomial Tables

n = 10									
x	…	π=.20	π=.25	π=.30	**π=.35**	π=.40	π=.45	π=.50	
0	…	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010	10
1	…	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098	9
2	…	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439	8
3	…	0.2013	0.2503	0.2668	**0.2522**	0.2150	0.1665	0.1172	7
4	…	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051	6
5	…	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461	5
6	…	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051	4
7	…	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172	3
8	…	0.0001	**0.0004**	0.0014	0.0043	0.0106	0.0229	0.0439	**2**
9	…	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098	1
10	…	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0
	…	π=.80	**π=.75**	π=.70	π=.65	π=.60	π=.55	π=.50	**x**

Examples:

n = 10, π = .35, x = 3: $P(x = 3|n = 10, \pi = .35) = .2522$

n = 10, π = .75, x = 2: $P(x = 2|n = 10, \pi = .75) = .0004$

POISSON PROBABILITY DISTRIBUTION

The Poisson Distribution Definitions

- You use the **Poisson distribution** when you are interested in the number of times an event occurs in a given **area of opportunity**.
- An **area of opportunity** is a continuous unit or interval of time, volume, or such area in which more than one occurrence of an event can occur.
  - The number of scratches in a car's paint
  - The number of mosquito bites on a person
  - The number of computer crashes in a day

The Poisson Distribution

- Apply the Poisson Distribution when:
  - You wish to count the number of times an event occurs in a given area of opportunity
  - The probability that an event occurs in one area of opportunity is the same for all areas of opportunity
  - The number of events that occur in one area of opportunity is independent of the number of events that occur in the other areas of opportunity
  - The probability that two or more events occur in an area of opportunity approaches zero as the area of opportunity becomes smaller
  - The average number of events per unit is λ (lambda)

Poisson Distribution Formula

$$P(X) = \frac{e^{-\lambda}\lambda^{x}}{X!}$$

where:

X = number of events in an area of opportunity

λ = expected number of events

e = base of the natural logarithm system (2.71828...)

Example

Assume that the number of network errors experienced in a day on a local area network (LAN) is distributed as a Poisson random variable. The mean number of network errors experienced in a day is 2.4. What is the probability that in any given day

a. zero network errors will occur?
b. exactly one network error will occur?
c. two or more network errors will occur?
d. fewer than three network errors will occur?

Example (Answer)

$\lambda = 2.4$

a. $P(X=0) = 0.0907$
b. $P(X=1) = 0.2177$
c. $P(X\geq 2)$
$$\begin{aligned} &= 1 - P(X<2) \\ &= 1 - (0.0907+0.2177) \\ &= 0.6916 \end{aligned}$$
d. $P(X<3)$
$$\begin{aligned} &= P(X=0) + P(X=1) + P(X=2) \\ &= 0.0907 + 0.2177 + 0.2613 \\ &= 0.5697 \end{aligned}$$

Poisson Distribution Characteristics

- Mean

$$\boxed{\mu = \lambda}$$

- Variance and Standard Deviation

$$\boxed{\sigma^2 = \lambda}$$

$$\boxed{\sigma = \sqrt{\lambda}}$$

where λ = expected number of events

Example

For the problem in the example before, what are the mean and standard deviation of the binomial distribution used in (a) through (d)? Interpret these values.

Using Poisson Tables

X	λ								
	0.10	0.20	0.30	0.40	**0.50**	0.60	0.70	0.80	0.90
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659
2	0.0045	0.0164	0.0333	0.0536	**0.0758**	0.0988	0.1217	0.1438	0.1647
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
4	0.0000	0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
5	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Example: Find $P(X = 2)$ if $\lambda = 0.50$

$$P(X=2) = \frac{e^{-\lambda}\lambda^{X}}{X!} = \frac{e^{-0.50}(0.50)^2}{2!} = 0.0758$$

EXERCISE

Exercise 1

The U.S. Department of Transportation reported that in 2009, Southwest led all domestic airlines in on-time arrivals for domestic flights, with a rate of 0.85. Using the binomial distribution, what is the probability that in the next six flights

a. four flights will be on time?
b. all six flights will be on time?
c. at least four flights will be on time?
d. What are the mean and standard deviation of the number of on-time arrivals?

Exercise 2

The quality control manager of Marilyn's Cookies is inspecting a batch of chocolate-chip cookies that has just been baked. If the production process is in control, the mean number of chip parts per cookie is 6.0. What is the probability that in any particular cookie being inspected

a. fewer than five chip parts will be found?
b. exactly five chip parts will be found?
c. five or more chip parts will be found?
d. either four or five chip parts will be found?

CLASS EXERCISE

1

It has been observed that the average number of traffic accidents requiring medical assistance on the Hollywood Freeway between 7 and 8 AM on Wednesday mornings is 1. What, then, is the chance that there will be a need for exactly 2 ambulances on the Freeway, during that time slot on any given Wednesday morning? The hospital dispatcher needs to know.

2

The probability that a student is accepted to a prestigious college is 0.3. If 5 students from the same school apply, what is the probability that at most 2 are accepted?

3

Entry to a certain University is determined by a national test. The scores on this test are normally distributed with a mean of 500 and a standard deviation of 100. Tom wants to be admitted to this university and he knows that he must score better than at least 70% of the students who took the test. Tom takes the test and scores 585. Will he be admitted to this university?

3

If Tim, a friend of Tom, scores 534, will he be admitted to this university

4

A bottling company uses a machine to fill the bottles with olive oil. The bottles are designed to contain 475 milliliters (ml). In fact, the contents vary according to a normal distribution with a mean of 473 ml and standard deviation of 3 ml.

a. What is the distribution, mean, and standard error of the sample mean of six randomly selected bottles?
b. What is the probability that the mean of six bottles is less than 470 ml?
c. What is probability that the mean of six bottles is more than 475 ml?

5

Identical computer components are shipped in boxes of 5. About 15% of components have defects. Boxes are tested in a random order.

a. What is the probability that a randomly selected box has only non-defective components?
b. What is the probability that at least 8 of randomly selected 10 boxes have only non-defective components?

6

Coliform bacteria are randomly distributed in a certain Arizona river at an average concentration of 1 per 20cc of water. If we draw from the river a test tube containing 10cc of water, what is the chance that the sample contains exactly 2 coliform bacteria?

7

A radar unit is used to measure speeds of cars on a motorway. The speeds are normally distributed with a mean of 90 km/hr and a standard deviation of 10 km/hr. What is the probability that a car picked at random is travelling at more than 100 km/hr?

8

The switchboard in a Denver law office gets an average of 2.5 incoming phone calls during the noon hour on Thursdays. Experience shows that the existing lunch hour staff can handle up to 5 calls in an hour. They seem to be well covered. But just to test the edges, what is the actual chance that 6 calls will be received during the lunch period, on some particular Thursday?

9

A large group of students took a test in Physics and the final grades have a mean of 70 and a standard deviation of 10. If we can approximate the distribution of these grades by a normal distribution, what percent of the students

a. scored higher than 80?
b. should pass the test (grades≥60)?
c. should fail the test (grades<60)?

10

The engines made by Ford for speedboats had an average power of 220 horsepower (HP) and standard deviation of 15 HP.

a. A potential buyer intends to take a sample of four engines and will not place an order if the sample mean is less than 215 HP. What is the probability that the buyer will not place an order?
b. If the customer samples 100 engines, what is the probability that the sample mean will be less than 215?

11

The number of calls coming per minute into a hotels reservation center is Poisson random variable with mean 3.

a. Find the probability that no calls come in a given 1 minute period.
b. Assume that the number of calls arriving in two different minutes are independent. Find the probability that atleast two calls will arrive in a given two minute period.

12

The length of similar components produced by a company are approximated by a normal distribution model with a mean of 5 cm and a standard deviation of 0.02 cm. If a component is chosen at random

a. what is the probability that the length of this component is between 4.98 and 5.02 cm?
b. what is the probability that the length of this component is between 4.96 and 5.04 cm?

13

A (blindfolded) marksman finds that on the average he hits the target 4 times out of 5. If he fires 4 shots, what is the probability of

a. more than 2 hits?
b. at least 3 misses?

14

A group of 625 students has a mean age of 15.8 years with a standard deviation of 0.6 years. The ages are normally distributed.

a. How many students are younger than 16.2 years? Express answer to the nearest student?

b. How many students are older than 14 years? Express answer to the nearest student?

15

A quality control engineer is in charge of testing whether or not 90% of the DVD players produced by his company conform to specifications. To do this, the engineer randomly selects a batch of 12 DVD players from each day's production. The day's production is acceptable provided no more than 1 DVD player fails to meet specifications. Otherwise, the entire day's production has to be tested.

a. What is the probability that the engineer incorrectly passes a day's production as acceptable if only 80% of the day's DVD players actually conform to specification?
b. What is the probability that the engineer unnecessarily requires the entire day's production to be tested if in fact 90% of the DVD players conform to specifications?

16

Penn State Fleet which operates and manages car rentals for Penn State employees found that the tire lifetime for their vehicles has a mean of 50,000 miles and standard deviation of 3500 miles.

a. What would be the distribution, mean and standard error mean lifetime of a random sample of 50 vehicles?
b. What is the probability that the sample mean lifetime for these 50 vehicles exceeds 52,000?

THANK YOU