MAT 1372 Stat w/ Prob classwk 21 Fall 2016

**7.3 SAMPLE MEAN**

Before we jump to the new material, let’s review some important definitions:

**Expectation:** (for discrete random variables)

**Variance:**

and properties:





 : provided that X and Y are independent!



Suppose that we sample n times from a population with mean μ and std dev σ. The **sample mean** is defined to be



Exercise: use the properties above to prove

1. 
2. 

For result 2 above, if we take the square root of both sides, we get



If theare the standard normal distribution, some of the sample mean distributions are plotted below:



Perhaps surprisingly, the sample means for non-normal distributions become more and more normal as n gets large, a result known as the **CENTRAL LIMIT THEOREM**.

**1.** Consider the population consisting of equal numbers of 1’s and 2’s. Plot the possible values along with their probabilities of the sample mean of a sample of specified size

*n* = 2 is done in the book as Example 7.1



**(a)** *n* = 3 (done in class, below)

**(b)** *n* = 4 (for homework)

Also derive the variance and standard deviation of the sample mean.

Are your answers consistent with the formulas presented in this section?

Clearly, μ = 1.5. To find variance, we use 



If flip a coin 3 times, there are 8 possible outcomes, each with equal probability:

111,112,121,122,211,212,221,222

We can group them base on their sums:

3 111

4 112,121, 211

5 122, 212,221

6 222

Remember, that in addition to summing, we want to divide by n, which is 3

The result is



which is consistent with the result 

To find the variance, we use the property



(the book uses the definition)





So 

which is consistent with the result 

**2.** Suppose that *X*1 and *X*2 constitute a sample of size 2 from a population

in which a typical value *X* is equal to either 1 or 2 with respective probabilities

*P*{*X* = 1} = 0.7 *P*{*X* = 2} = 0.3

**(a)** Compute *E*[*X*].

**(b)** Compute Var(*X*).

**(c)** What are the possible values of = (*X*1 + *X*2) /2?

**(d)** Determine the probabilities that assumes the values in (c).

**(e)** Using (d), directly compute and 

**(f)** Are your answers to (a), (b), and (e) consistent with the formulas

presented in this section?

Solution:

 **(a)** 

**(b)** 

**(c), (d)**



**(e), (f)** 

which are in agreement with and 

 **4.** The amount of money withdrawn in each transaction at an automatic

teller of a branch of the Bank of America has mean $80 and standard

deviation $40. What are the mean and standard deviation of the average

amount withdrawn in the next 20 transactions?

**7.** The weight of a randomly chosen person riding a ferry has expected

value 155 pounds and standard deviation 28 pounds. The ferry has

the capacity to carry 100 riders. Find the expected value and standard

deviation of the total passenger weight load of a ferry at capacity.

**7.4 CENTRAL LIMIT THEOREM**

Given , a family of RV’s each with mean μ and SD σ, define



then

  

Exercise: Prove results for using the fact that , the corresponding results for and the properties of expectation and SD reviewed last class.

**Central Limit Theorem** For any  with meanand SD σ,

1. has mean and SD ;
2. has mean and SD  ;
3. and  become well-approximated by normal distributions for larger n

To the right above is an example from text: exponential distribution .

How large does n have to be for and to be “normal”. For most purposes n=30 is sufficient.

For an animation using a uniform distribution for see

<http://www.statisticalengineering.com/images/CLTuniform.gif>

When doing your homework exercises, you should first identify which distribution you will use, i.e., or . Reminder that no credit will be given unless accompanied by a normal graph with both standard and nonstandard horizontal labelings and with the appropriate inequality shaded.

 I illustrate one of each, and then have students work in groups on additional problems. At the end of the period, each group will present its problem.

**2.** Frequent fliers of a particular airline fly a random number of miles

each year, having mean and standard deviation (in thousands of miles)

of 23 and 11, respectively. As a promotional gimmick, the airline has

decided to randomly select 20 of these fliers and give them, as a bonus,

a check of $10 for each 1000 miles flown. Approximate the probability

that the total amount paid out is

**(a)** Between $4500 and $5000

**(b)** More than $5200

The trick to deciding whether to use or  is to determine whether an average or a sum of outcomes is being asked for. We are asked for a total not an average payout so we want to work with.

μ and σ are given in thousands of miles. To convert to $, multiply by 10 (payouts are $10 for each thousand miles) to get μ=$230 and σ=$110. Hence,

  and 

1. Form the appropriate inequality:



and standardize





Using Excel,

=NORM.S.DIST(4\*SQRT(5)/11,true)-NORM.S.DIST(-SQRT(5)/11,true)=.3725

Finish with a sentence:

The probability that the total payout is between $4500 and $5000 is 37%.

1. Similarly



and standardize





Using Excel,

=1-NORM.S.DIST(6\*SQRT(5)/11,true)=.1113

The probability that the total payout is over $5200 is 11%.

**14.** The lifetime of a certain type of electric bulb has expected value 500

 hours and standard deviation 60 hours. Approximate the probability

that the sample mean of 20 such light bulbs is less than 480 hours.

Here we are asked for an average (sample mean) so we work with:

 and 

The problem is to find the probability that .

Standardize





Using Excel: =NORMSDIST(-2\*SQRT(5)/3)=.06802

The probability that the sample mean of 20 bulbs is less than 480 hrs is 6.8%.

Now it’s your turn:

 Working in groups of 2 or 3, select a problem 4, 6, 8, 9, 10, 12, 15 or 16.

**4.** If you place a $1 bet on a number of a roulette wheel, then either you

win $35, with probability 1/38, or you lose $1, with probability 37/38.

Let *X* denote your gain on a bet of this type.

**(a)** Find *E*[*X*] and SD(*X*).

Suppose you continually place bets of the preceding type. Show that

**(b)** The probability that you will be winning after 1000 bets is approximately

0.39.

**(c)** The probability that you will be winning after 100,000 bets is

approximately 0.002.

**6.** A zircon semiconductor is critical to the operation of a superconductor

and must be immediately replaced upon failure. Its expected

lifetime is 100 hours, and its standard deviation is 34 hours. If 22

of these semiconductors are available, approximate the probability

that the superconductor can operate for the next 2000 hours. (That is,

approximate the probability that the sum of the 22 lifetimes exceeds

2000.)

**8.** A highway department has enough salt to handle a total of 80 inches of

snowfall. Suppose the daily amount of snow has a mean of 1.5 inches

and a standard deviation of 0.3 inches.

**(a)** Approximate the probability that the salt on hand will suffice for

the next 50 days.

**(b)** What assumption did you make in solving part (a)?

**(c)** Do you think this assumption is justified? Explain briefly!

**9.** Fifty numbers are rounded off to the nearest integer and then summed.

If the individual roundoff errors are uniformly distributed between

–0.5 and 0.5, what is the approximate probability that the resultant

sum differs from the exact sum by more than 3? (Use the fact that the

mean and variance of a random variable that is uniformly distributed

between –0.5 and 0.5 are 0 and 1/12, respectively.)

**10.** A six-sided die, in which each side is equally likely to appear, is

repeatedly rolled until the total of all rolls exceeds 400. What is the

approximate probability that this will require more than 140 rolls?

(*Hint*: Relate this to the probability that the sum of the first 140 rolls is

less than 400.)

**12.** Suppose that the number of miles that an electric car battery functions

has mean μ and standard deviation 100. Using the central limit

theorem, approximate the probability that the average number of miles

per battery obtained from a set of *n* batteries will differ from μ by more

than 20 if

1. *n* = 10 **(b)** *n* = 20 **(c)** *n* = 40 **(d)** *n* = 100

**15.** Consider a sample of size 16 from a population having mean 100 and

standard deviation σ. Approximate the probability that the sample

mean lies between 96 and 104 when

1. σ = 16 **(b)** σ = 8 **(c)** σ = 4 **(d)** σ = 2 **(e)** σ = 1

**16.** An instructor knows from past experience that student examination

scores have mean 77 and standard deviation 15. At present, the

instructor is teaching two separate classes—one of size 25 and the

other of size 64.

**(a)** Approximate the probability that the average test score in the

class of size 25 lies between 72 and 82.

**(b)** Repeat (a) for the class of size 64.

**(c)** What is the approximate probability that the average test score in

the class of size 25 is higher than that in the class of size 64?

**(d)** Suppose the average scores in the two classes are 76 and 83.

Which class—the one of size 25 or the one of size 64—do you think

was more likely to have averaged 83? Explain your intuition.