AP Statistics Name Key
Unit 04 – Probability Period
Day 06 Notes – Transforming & Combining Random Variables

General Rules for Means and Variances

1. Adding or subtracting a constant from data shifts the mean (and other measures of center: median, quartiles, percentiles) but doesn't change the variance or standard deviation (or other measures of spread: range, IQR). It does not change the shape of the distribution.

$$E(X \pm c) = E(X) \pm c \qquad Var(X \pm c) = Var(X)$$

$$SD(X \pm c) = SD(X)$$

EXAMPLE:

Suppose that for several weeks the restaurant has also been distributing coupon worth $5 off any one meal (one discount per table). If every couple dining there on Valentine's Day brings a coupon, what will be the mean and standard deviation of the total discounts they'll receive?
Remember that the mean discount was expected to be $5.83 with a standard deviation of $8.62.

Let D = total discount (Lucky Lovers plus the coupon);
Then D = X + 5

E(D) = E(X+5) = E(X) + 5 = $5.83 + 5 = \$10.83$

Var(D) = Var(X+5) = Var(X) = $8.62^2 = \$74.30$

SD(D) = $8.62 ← doesn't change

2. Multiplying (and dividing by a constant) each value of a random variable by a constant multiplies the mean (and other measures of center: median, quartiles, percentiles) by that constant and the variance by the square of the constant. It also multiples other measures of spread (range, IQR, SD) by that constant. It does not change the shape of the distribution.

$$E(aX) = a \cdot E(X) \qquad Var(aX) = a^2 \cdot Var(X) \qquad SD(aX) = a \cdot SD(X)$$

EXAMPLE:

When two couples dine together on a single check, the restaurant doubles the discount offer ($40 for finding the ace of hearts on the first try and $20 for the second try). What are the mean and standard deviation of discounts for such foursomes?

$E(2X) = 2E(X) = 2 \cdot 5.83 = \$11.66$

$Var(2X) = 2^2 \cdot Var(X) = 4 \cdot 8.62^2 = \$297.22$

$SD(2X) = 2 \cdot SD(X) = 2 \cdot 8.62 = \$17.24$

3. Linear Transformations: a combination of adding/subtracting and multiplying/dividing numbers to the random variable.

If Y = a + bX is a linear transformation of the random variable X, then…

- The probability distribution of Y has the same shape as the probability distribution of X if b > 0.

- $\mu_Y = a + b\mu_X$

- $\sigma_Y = |b|\,\sigma_X$ (because b could be a negative number)

EXAMPLE:

One brand of bathtub comes with a dial to set the water temperature. When the "babysafe" setting is selected and the tub is filled, the temperature X of the water follows a Normal distribution with a mean of 34 degrees C and a standard deviation of 2 degrees C.

$N(34, 2)$

Define the random variable Y to be the water temperature in degrees Fahrenheit when the dial is set on "babysafe." Recall that F = (9/5)C + 32. Find the mean and standard deviation of Y.

this is the linear transformation

we must convert °C to °F.

$Y = \frac{9}{5}X + 32$

$\mu_Y = 32 + \frac{9}{5}\mu_X$ 34 $= 32 + \frac{9}{5}(34) = 93.2°F$

$\sigma_Y = \left|\frac{9}{5}\right|\sigma_X = \left|\frac{9}{5}\right| \cdot 2 = 3.6°F$

when the dial is set on "babysafe," the mean water temperature is 93.2°F with a standard deviation of 3.6°F.

According to Babies R Us, the temperature of a baby's bathwater should be between 90 degrees Fahrenheit and 100 degrees Fahrenheit. Find the probability that the water temperature on a randomly selected day when the "babysafe" setting is used meets the Babies R Us recommendation. Show your work.

$N(93.2, 3.6)$ we know

$P(90 \le Y \le 100)$? we want to find

$z = \frac{x-\mu}{\sigma} = \frac{90-93.2}{3.6} = -0.89$ and $z = \frac{100-93.2}{3.6} = 1.89$

$P(-0.89 < z < 1.89) =$ There is 78.4% chance that the water temperature on a randomly selected day when the "babysafe" setting is used meets the Babies R Us recommendation.

Independent Random Variables: If knowing whether any event involving X alone has occurred tells us nothing about the occurrence of any event involving Y alone, and vice versa, then X and Y are independent random variables.

4. Combining Random Variables: Adding or subtracting events.

$E(X \pm Y) = E(X) \pm E(Y)$ $Var(X \pm Y) = Var(X) + Var(Y)$

EXAMPLE:

You and your partner go to this restaurant with another couple and agree to share any benefit from this promotion. Does it matter whether you pay separately or together?

Yes! When couples pay with separate checks, there is less variation in their total discount. The standard deviation is \$12.19, compared to \$17.24 for couples who pay for the double discount on a single check.

X_1 = your account
X_2 = your friends' account

T = total

$E(T) = E(X_1 + X_2) = E(X_1) + E(X_2) = \$5.83 + \$5.83 = \$11.66$

$Var(T) = Var(X_1 + X_2) = Var(X_1) + Var(X_2) = \$8.62^2 + \$8.62^2 = \$148.608$

$SD(T) = \sqrt{148.608} = \$12.19$

5. Combining Random Variables: Summing a series of outcomes.

EXAMPLE:

The restaurant owner is planning to serve 40 couples on Valentine's Day. What's the expected total of discounts the owner will give? With what standard deviation?

$E(T) = E(X_1 + X_2 + X_3 + \ldots + X_{39} + X_{40})$

$E(T) = 5.83 + 5.83 + 5.83 + \ldots + 5.83 + 5.83 = 5.83 \cdot 40$

$E(T) = \$\ 233.20$

$SD(T) = \sqrt{Var(x_1 + x_2 + x_3 + \cdots .. x_{39} + x_{40})}$

$SD(T) = \sqrt{8.62^2 + 8.62^2 + 8.62^2 + \ldots + 8.62^2 + 8.62^2} = \$54.52$

The restaurant owner can expect the 40 couples to win discounts totaling \$233.20 with a standard deviation of \$54.52.

6. Normal Distribution: The sum or difference of independent Normal random variables follows a Normal distribution.

EXAMPLE:

Consider a company that manufactures and ships small stereo systems. The times required to pack the stereos can be described by a Normal model with a mean of 9 minutes and a standard deviation of 1.5 minutes. The times for the boxing stage can also be modeled as Normal, with a mean of 6 minutes and a standard deviation of 1 minute.

What is the probability that packing two consecutive systems takes over 20 minutes?

1. State the problem
   $P_{1=}$ time for packing the first system
   P_2 = time for packing the second
   T = total time to pack two systems
   $T = P_1 + P_2$

2. Find the expected value and the standard deviation.
   $E(T) = E(P_1 + P_2)$
   $E(T) = E(P_1) + E(P_2)$
   $E(T) = 9 + 9 = 18$ minutes

   $Var(T) = Var(P_1 + P_2)$
   $Var(T) = Var(P_1) + Var(P_2)$
   $Var(T) = 1.5^2 + 1.5^2$
   $Var(T) = 4.50$
   SD(T)= $\sqrt{4.50} \approx 2.12$ minutes

3. Since we are using a Normal distribution you should use the Normal bell curve z-score. The mean is 18 and you want to find it compared to 20 minutes so…

$$z = \frac{20-18}{2.12} = 0.94$$

Now take that number and look on the z-table to find the probability = 17.36%

4. Interpret your result in context.
There's a little more than 17.36% chance that it will take a total of over 20 minutes to pack two consecutive stereo systems.

What percentage of the stereo systems takes longer to pack than to box?

1. Define your random variables
   P = time for packing a system
   B = time for boxing a system
   D = difference in times to pack and box a system

   D = P – B

2. Find the expected value and the standard deviation
   $E(D) =$ E(P−B)
   $E(D) =$ E(P) − E(B)
   $E(D) =$ 9−6 = 3 minutes
   $Var(D) =$ Var(P−B)
   $Var(D) =$ Var(P) + Var(B)
   $Var(D) =$ $1.5^2 + 1^2 = 3.25$
   $SD(D) =$ $\sqrt{3.25} = 1.8$ minutes

3. Find the z-score value for the Normal distribution

$$z = \frac{0-3}{1.8} = -1.67$$

Look up this number in the z-table and you get 0.9525

4. Interpret your result in context.
   About 95.25% of all the stereo systems will require more time for packing than for boxing.

Practice problems:

1. A delivery company's trucks occasionally get parking tickets, and based on past experience, the company plans that the trucks will average 1.3 tickets a month, with a standard deviation of 0.7 tickets.

   If they have 18 trucks, what are the mean and standard deviation of the total number of parking tickets the company will have to pay this month?

   $\mu = 1.3$ tickets
   $\sigma = 0.7$ tickets

   $E(T) = 18 \cdot 1.3 = 23.4$ tickets
   $Var(T) = 18^2 \cdot 0.7^2 = 158.76$
   $SD(T) = \sqrt{158.76} \approx 12.6$ tickets

   The company can expect to pay 23.4 tickets with a standard deviation of 12.6 tickets.

2. The amount of cereal that can be poured into a small bowl varies with a mean of 1.5 ounces and a standard deviation of 0.3 ounces. A large bowl holds a mean of 2.5 ounces with a standard deviation of 0.4 ounces. You open a new box of cereal and pour one large and one small bowl. L = large bowl S = small bowl T = total = L + S

   a. How much more cereal do you expect to be in the large bowl?

   $E(L-S) = E(L) - E(S) = 2.5 - 1.5 = 1$ ounce more cereal is expected to be in the large bowl than the small bowl.

   b. What's the standard deviation of this difference?

   $Var(L-S) = Var(L) + Var(S) = 0.4^2 + 0.3^2 = 0.25$
   $SD(L-S) = \sqrt{0.25} \approx$ 0.5 ounces

c. What are the mean and standard deviation of the total amount of cereal in the two bowls?

$E(T) = E(S+L) = E(S) + E(L) = 1.5 + 2.5 = 4$ ounces

$var(S+L) = var(S) + var(L) = 0.5$ ounces

d. If the total follows a Normal model, what's the probability you poured out more than 4.5 ounces of cereal in the two bowls together? $N(4, 0.5)$

$z = \frac{x-\mu}{\sigma} = \frac{4.5-4}{0.5} = 1 \rightarrow P(z>1) = 15.87\%$

There is a 15.87% chance that you poured out more than 4.5oz of cereal in the 2 bowls together.

e. The amount of cereal the manufacturer puts in the boxes is a random variable with a mean of 16.3 ounces and a standard deviation of 0.2 ounces. Find the expected amount of cereal left in the box and the standard deviation.

$E(box - T) = E(box) - E(T) = 16.3 - 4 = 12.3$ oz

$var(box - T) = var(box) + var(T) = 0.2^2 + 0.5^2 = 0.29$

$SD(box - T) = \sqrt{0.29} \approx 0.54$ oz

The expected amount of cereal left in the box is 12.3oz w/ an SD of 0.54oz.

3. A time-and-motion study measures the time required for an assembly-line worker to perform a repetitive task. The data show that the time required to bring a part from a bin to its position on an automobile chassis varies from car to car according to a Normal distribution with mean 11 seconds and standard deviation 2 seconds. The time required to attach the part of the chassis follows a Normal distribution with mean 20 seconds and standard deviation 4 seconds. The study finds that the times required for the two steps are independent. A part that takes a long time to position, for example, does not take more or less time to attach than other parts.

$N(11, 2)$ $N(20, 4)$

a. What is the distribution of the time required for the entire operation of position and attaching a randomly selected part?

T = total time = move + attach

$E(T) = E(move + attach) = E(move) + E(attach) = 11 + 20 = 31$ seconds

$var(T) = 2^2 + 4^2 = 20$

$SD(T) = \sqrt{20} \approx 4.47$ seconds

The time required for the entire operation is 31 seconds w/ an SD of 4.47 seconds

b. Management's goal is for the entire process to take less than 30 seconds. Find the probability that this goal will be met for a randomly selected part.

$z = \frac{x-\mu}{\sigma} = \frac{30-31}{4.47} = -0.22$

$P(z < -0.22) = 41.29\%$

The probability that the entire process will take less than 30 seconds is 41.29% for a randomly selected part. for a randomly selected part.