EFFECTS OF A LINEAR TRANSFORMATION ON A RANDOM VARIABLE

The final calculation of Pete's Jeep Tour profit can be given by the equation $V=150x-100$, or equivalently $V= -100 + 150x$. This linear transformation uses both multiplication/division and addition/subtraction.

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 is b>0
- mu(y) = a + bmu(x)
- Sigma(y)=abs(b)sigma(x) (because b could be a negative number)

LINEAR TRANSFORMATIONS, CONT'D

The Baby and the Bathwater

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 Celsius and a standard deviation of 2 degrees Celsius.

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

b) According to Babies R Us, the temperature of a baby's bathwater should be between 90 degrees 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.

a) Mu(y) = 32 + 9/5(34) = 93.2 degrees F ; SD = 9/5 (2) = 3.6 degrees F
b) Transforming does not change shape, so this is still a normal distribution. Draw a histogram and shade in the desired area. Calculate z scores and find what's between:

z = 90 – 93.2/3.6 = -.89
p = .1867

z = 100-93.2/3.6 = 1.89
p = .9706

$P(90 \leq Y \leq 100) = .9706 - .1867 = 0.7839$.

There's about a 78% chance that the water temperature meets the recommendation on a randomly selected day.

COMBINING RANDOM VARIABLES

Back to Pete's Jeep Tours again – Pete's sister Erin is impressed by the success of Pete's business. She decides to join the business, running tours on the same days as Pete in her slightly smaller vehicle under the name "Erin's Adventures." After a year of steady bookings, Erin discovers that the number of passengers Y on her tours have the following probability distribution:

No. of passengers y_i	**2**	**3**	**4**	**5**
Probability p_i	0.3	0.4	0.2	0.1

Mean: 3.10 Stand. Dev. 0.943

Recall Pete's distribution:

No. of passengers x_i	**2**	**3**	**4**	**5**	**6**
Probability p_i	0.15	0.25	0.35	0.20	0.05

Mean: 3.75 Stand. Dev. 1.0897

How many total passengers *T* will Pete and Erin have on their tours on a randomly selected day?

How many more or fewer passengers, *D*, will Pete have than Erin on a randomly selected day?

Just add up their means!
3.10 + 3.75 = 6.85 on average on a randomly selected day
T = X + Y

Add their ranges for variability:
D = X – Y
D = Range of T = Range of X + Range of Y ➔ (6-2)+(5-2) = 4 + 3 = 7
Or think of it as : Total possible passengers – fewest possible passengers = 11 - 4 = 7
(remember, more variables mean more variability)

What about standard deviation? The events MUST be independent so we can multiply them! Otherwise, we're stuck.
Ex. T = 4; only happens when Pete has 2 (p=0.15) and Erin has 2 (p=0.3). Luckily, knowing one does NOT affect us knowing the other, so P(T=4) = .045

COMBINING RANDOM VARIABLES – STANDARD DEVIATION

Let T = X+Y, as before. Because X and Y are independent random variables, we can construct the probability distribution by listing all combinations of X and Y that yield each possible value of T and adding the corresponding probabilities.

Value t_i	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Probability p_i								

Ex: P(T=4) = P(X=4) • P(Y=4) = .15 x .3 = .045
P(.T=5) = P(X=3) • P(Y=2) + P(X=2) •P(Y=3) = .135
P(T=6) = x=4•y=2 + y=4•x=2 = .235
P(T=7) = .265
P(T=8) = .190
P(T=9) = .095
P(T=10)= .030
P(T=11)= .005

Mean = 4(.045) + 5(.135)+...+11(.005) = 6.85
**this is also MuX + MuY

ST = $(4-6.85)^2(.045)$ +...+$(11-6.85)^2(.005)$ = 2.0775 → sqrt = 1.441
**the variance of this data set (2.0775) is just VarX + VarY; you can add variances and then square root for stnd deviation

COMBINING RANDOM VARIABLES: SYNTHESIS

- Mean of the sum of Random Variables
- Mean of the difference of Random Variables
- Variance of the Sum (or difference) of Independent Random Variables

Mean of Sum: For any two random variables X and Y, if T = X + Y, then the expected value of T is E(T) = mu_t = mu_{x+y} = mu_x + mu_y ; in general, the mean of the sum of several random variables is the sum of their means

Mean of difference: D = X-Y then mu_D = mux – mu_y --> order of subtraction DOES MATTER! Pay attention to your question to be sure you're in the right order!

Variance: For any two *independent* random variables X and Y, if T = X+Y, then the variance of T is var_t = var_x + var_y; standard deviation of T is sqrt(var_t) (DO NOT ADD STANDARD DEVIATIONS)

** make sure the events are independent: 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

*Even with difference, more variables means more variability

EXAMPLE: ADDING RANDOM VARIABLES

Earlier, we defined X = the number of passengers that Pete has and Y = the number of passengers that Erin has on a randomly selected day. Recall that:

$\mu_x = 3.75, \sigma_x = 1.0897$ $\mu_y = 3.10, \sigma_y = 0.943$

Pete charges $150 per passenger and Erin charges $175 per passenger.

Calculate the mean and standard deviation of the total amount that Pete and Erin collect on a randomly chosen day.

μ_T= (3.75)(150) + 3.10(175) = 1105

Because the number of passengers are independent,

$(150)\sigma_x = 163.46$ and $(175)\sigma_y = 165.03$
$\sigma_T = (163.46)^2 + (165.03)^2 = 53{,}954.07$ → sqrt = 232.28

On average, Pete and Erin expect to collect $1105 per day, with a standard deviation of $232.28

EXAMPLE: DIFFERENCE OF RANDOM VARIABLES

Calculate the mean and standard deviation of the difference in the amounts that Pete and Erin collect on a randomly chosen day. Interpret each value in context.

D = X – Y

$\mu_D = 150\mu_X - 175\mu_Y$ = 562.50 – 542.50 = 20.00

On average, Pete collects $20 more per day than Erin does.

σ_D = 232.28 → same as before because variance always adds!

CHECK YOUR UNDERSTANDING

A large auto dealership keeps track of sales and lease agreements made during each hour of the day. Let X = the number of cars sold and Y = the number of cars leased during the first hour of business on a randomly selected Friday. Based on previous records, the probability distributions of X and Y are as follows:

Cars sold x_i:	0	1	2	3
Probability p_i:	0.3	0.4	0.2	0.1

Mean: $\mu_X = 1.1$ Standard deviation: $\sigma_X = 0.943$

Cars leased y_i:	0	1	2
Probability p_i:	0.4	0.5	0.1

Mean: $\mu_Y = 0.7$ Standard deviation: $\sigma_Y = 0.64$

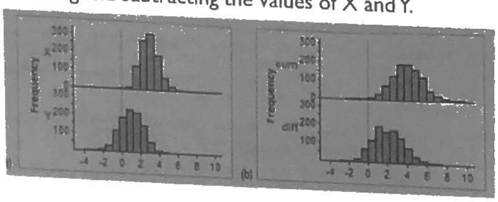
Define $D = X - Y$. Assume that X and Y are independent.

1. Find and interpret μ_D.
2. Compute σ_D. Show your work.
3. The dealership's manager receives a \$500 bonus for each car sold and a \$300 bonus for each car leased. Find the mean and standard deviation of the difference in the manager's bonus for cars sold and leased. Show your work.

Homework!

COMBINING NORMAL RANDOM VARIABLES

Normally distributed random variables, X and Y, have distributions N(3, 9) and N(1, 1.2), respectively. What do we know about the sum and difference of these two random variables? The histograms come from adding and subtracting the values of X and Y.



Let's summarize what we see:

Sum X + Y
Shape: approx normal
Center: about 4 (which makes sense bc 3 + 1 = 4)

Difference X – y
Shape: approx normal
Center: about 2 (which makes sense bc 3-1 =2)

Spread: spread looks the same for both, which makes sense because varx + vary = varx+y

**Any sum or difference of independent Normal random variables is also Normally distributed!

SUMS OF NORMAL RANDOM VARIABLES

Mr. Starnes likes sugar in his hot tea. From experience, he needs between 8.5 and 9 grams of sugar in a cup of tea for the drink to taste right. While making his tea one morning, he add four randomly selected packets of sugar. Suppose the amount of sugar in these packets follows a Normal distribution with mean 2.17 grams and standard deviation 0.08 grams.

What's the probability that Mr. Starnes' tea tastes right?

Step 1: State the distribution and values of interest – Let X = the amount of sugar in a randomly selected packet. Then X_i = amount of sugar in packet i (list them out). Each of these random variables has a normal distribution with mean 2.17 grams and standard deviation of .08 grams. We're interested in the total amount of sugar that Mr. Starnes puts in his tea, which is given by $T = X_1 + X_2 + X_3 + X_4$
Since T is the sum of four independent Normal random variables, mu(T) = Mu(x1) + mu(x2) + mu(x3) + mu(x4) = 2.17 + 2.17+2.17+2.17=8.68 grams,
And variance $(0.08)^2 + (0.08)^2 + (0.08)^2 + (0.08)^2 = .0256$ → sqrt = 0.16 grams.
We want to find the probability that the total amount of sugar in the tea is between 8.5 and 9 grams.
Draw density curve!

Step 2: Perform calculations
To find the area under the density curve, we standardize the boundary values using table A:
Z = 8.5-8.68/0.16 = -1.13 and z = 9 – 8.68/0.16 = 2
Then $P(-1.13 \leq Z \leq 2.00) = 0.9772-0.1292=0.8480$

Step 3: Answer the question

DIFFERENCES OF NORMAL RANDOM VARIABLES

The diameter C of a randomly selected large drink cup at a fast food restaurant follows a Normal distribution with a mean of 3.96 inches and a standard deviation of 0.01 inches. The diameter L of a randomly selected large lid at this restaurant follows a Normal distribution with mean 3.98 inches and standard deviation 0.02 inches. For a lid to fit a cup, the value of L has to be bigger than the value of C, but not by more than 0.06 inches.

Problem: What's the probability that a randomly selected large lid will fit on a randomly chosen large drink cup?

Solution:
Step 1: State the distribution and the values of interest
We'll define the random variable D = L – C to represent the difference between the lid's diameter and the cup's diameter. The random variable D is the difference of two independent Normal random variables. So D follows a Normal distribution with mean
Mu(d) = mu(L) – mu© = 3.98-3.96=.02
And variance $(0.02)^2 + (.01)^2 = .0005$ → sqrt = .0224

We want to find the probability that the difference D is between 0 and 0.06 inches. Draw a density curve and find shaded region here.

Step 2: We standardize the boundary values and use Table A:
Z = 0 - .02/.0224 = -.89 z = .06-.02/.0224 = 1.79
Then $P(-.89 \leq Z \leq 1.79) = .9633-.1867 = .7766$

Step 3: Answer the question:
There's about a 78% chance that a randomly chosen large lid will fit a randomly chosen large drink cup at this restaurant. Roughly 22% of the time, the lid won't fit.