

Probability and Statistics - Open and Free

[- \[SIGN IN \]](#)[Syllabus](#) | [Outline](#) | [Help](#) | [More](#)

Unit 3 :: Probability

This course is not led by an instructor

[Random Variables](#)[Sampling Distributions](#)

Module 8 / Mean and Variance of a Random Variable (5 of 5)

137

Find the mean and variance of a discrete random variable, and apply these concepts to solve real-world problems.

The concept of standard deviation is a bit harder to grasp than that of the mean. The purpose of the following examples and activities is to help you gain a better feel for the standard deviation of a random variable:

Xavier's and Yves' Production Lines

Recall the probability distribution of the random variable X , representing the number of defective parts per hour produced by Xavier's production line, and the probability distribution of the random variable Y , representing the number of defective parts per hour produced by Yves' production line:

X	0	1	2	3	4
$P(X=x)$.15	.30	.25	.20	.10

Y	0	1	2	3	4
$P(Y=y)$.05	.05	.10	.75	.05

Look carefully at both probability distributions. Both X and Y take the same possible values (0, 1, 2, 3, 4). However, they are very different in the way the probability is distributed among these values. We saw before that this makes a difference in means:

$$\mu_X = 1.8$$

$$\mu_Y = 2.7$$

We now want to get a sense about how the different probability distributions impact their standard deviations.

Recall that the standard deviation of a random variable can be interpreted as a typical (or the long-run average) distance between the value of X and its mean.

We found that the standard deviation of X is 1.21. Now look at both probability distributions carefully. Which of the following do you think is true about the standard deviation of Y ?

Hint

- ☐ The standard deviation of Y is greater than 1.21.
- ☐ The standard deviation of Y is less than 1.21.
- ☐ The standard deviation of Y is roughly equal to 1.21.

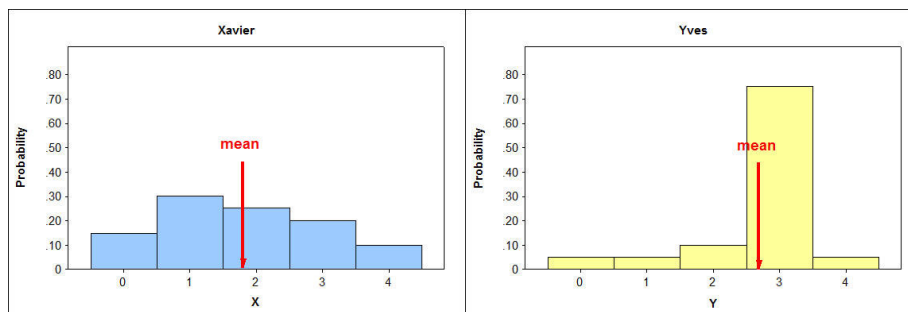
So, 75% of the time Y will assume a value (3) that is very close to its mean (2.7), while X will assume a value (2) that is close to its mean (1.8) much less often—only 25% of the time. The long-run average, then, of the distance between the values of Y and their mean will be much smaller than the long-run average of the distance between the values of X and their mean.

Therefore, $\sigma_Y < \sigma_X = 1.21$. Actually, $\sigma_Y = 0.85$, so we can draw the following conclusion:

Yves' production line produces an average of 2.70 defective parts per hour. The number of defective parts varies from hour to hour; typically (or, on average), it is about .85 away from 2.70.

Summary

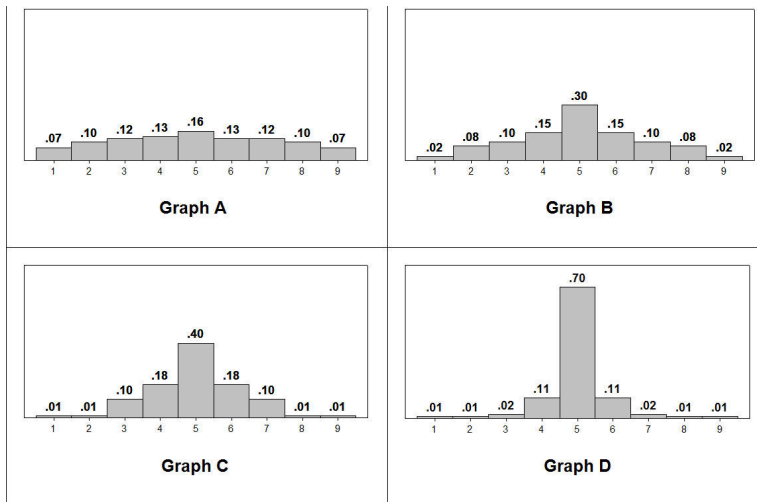
Here are the histograms for the production lines:



When we compare distributions, the distribution in which it is **more likely** to find values that are further from the mean will have a **larger** standard deviation. Likewise, the distribution in which it is **less likely** to find values that are further from the mean will have the **smaller** standard deviation.

In distribution, it is more likely to find values that are further from the mean.
Therefore, it is more likely that the standard deviation for this distribution is .

The following graphs will be used in the next "Did I Get This?" exercise.



The mean of each of the probability distributions shown above is 5. Which has the largest standard deviation?

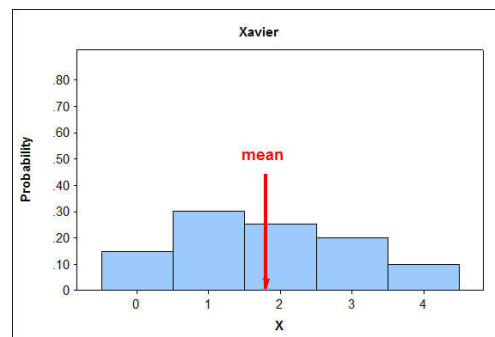
- ☐ Graph A
 ☐ Graph B
 ☐ Graph C
 ☐ Graph D

Comment

As we have stated before, using the mean and standard deviation gives us another way to assess which values of a random variable are unusual. Any values of a random variable that fall within 2 standard deviations of the mean would be considered ordinary (not unusual).

Xavier's Production Line—Unusual or Not?

Looking once again at the probability distribution for Xavier's production line:



Would it be considered unusual to have 4 defective parts per hour?

We know that $\mu_X = 1.8$ and $\sigma_X = 1.21$.

Ordinary values are within 2 standard deviations of the mean. $1.8 - 2(1.21) = -.62$ and $1.8 + 2(1.21) = 4.22$. This gives us an interval from $-.62$ to 4.22 . Since we cannot have a negative number of defective

parts, the interval is essentially from 0 to 4.22. Because 4 is within this interval, it would be considered ordinary. Therefore, it is **not unusual**.

Would it be considered unusual to have no defective parts? Zero is within 2 standard deviations of the mean, so it would not be considered unusual to have no defective parts.

The following activity will reinforce this idea.

Learn By Doing

Recall the probability distribution for changing majors.

We have made the following calculations for the mean and standard deviation. For some extra practice, feel free to verify our calculations.

$$\mu_X = 1.23 \text{ and } \sigma_X = 1.08$$

John's parents are concerned because John is changing majors for the second time. John claims this is not unusual. Using the mean and standard deviation given above, is John's behavior unusual?

"Risk" in investments provides a useful application for the concept of variability. If there is no variability at all in possible outcomes, then the outcome is something we can count on, with no risk involved. At the other extreme, if there is a large amount of variability with possibilities for either tremendous loss or gain, then the associated risk is quite high.

If a variable's possible values just differ somewhat, with some only marginally favorable and others unfavorable, then the underlying random experiment entails just a moderate amount of risk. The following example demonstrates how differing values of standard deviation reflect the amount of risk in a situation.

Comparing Investments

Consider three possible investments, with returns denoted as X, Y, and Z, respectively, and probability distributions outlined in the tables below.

X	14,000
P(X=x)	1

Investment X is what we'd call a "sure thing," with a guaranteed return of \$14,000: there is no risk involved at all.

Y	0	1,000,000
P(Y=y)	.98	.02

Investment Y is extremely risky, with a high probability (.98) of no gain at all, contrasted by a slight probability (.02) of "making a killing" with a return of a million dollars.

Z	10,000	20,000
P(Z=z)	.5	.5

Investment Z is somewhere in between: there is an equal chance for either a return that's on the low side or a return that's on the high side.

If you only consider the mean return on each investment, would you prefer X, Y, or Z? The means for X, Y, and Z are calculated as follows:

$$\mu_X = 14000(1) = 14000$$

$$\mu_Y = 0(0.98) + 1000000(0.02) = 20000$$

$$\mu_Z = 10000(0.5) + 20000(0.5) = 15000$$

Clearly, the mean return for Y is highest, and so investment in Y would seem to be preferable.

Now consider the standard deviations, and consider which investment you'd prefer—X, Y, or Z.

The standard deviations are:

$$\sigma_X^2 = (14000 - 14000)^2 (1) = 0$$

$$\sigma_X = 0$$

$$\sigma_Y^2 = (0 - 20000)^2 (0.98) + (1,000,000 - 20000)^2 (0.02) = 1.96 \times 10^{10}$$

$$\sigma_Y = 140,000$$

$$\sigma_Z^2 = (10000 - 15000)^2 (0.5) + (20000 - 15000)^2 (0.5) = 25,000,000$$

$$\sigma_Z = 5000$$

Granted, the mean returns suggest that investment X is least profitable and investment Y is most profitable. On the other hand, the standard deviations are telling us that the return for X is a sure thing; for Y, the remote chance of making a huge profit is offset by a high risk of losing the investment entirely; for Z, there is a modest amount of risk involved. If you can't afford to lose any money, then investment X would be the way to go. If you have enough assets to take a chance, then investment Y would be worthwhile. In particular, if a large company routinely makes many such investments, then in the long run there will occasionally be such enormous gains that the company is willing to absorb many smaller losses. Investment Z represents the middle ground, somewhere between the other two.

End of Lesson Questions

About the Mean and Variance of a Random Variable



Unless otherwise noted this work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#).

[Open Learning Initiative](#)

137