Section 7.3
Variance and Standard Deviation

The **Variance** of a random variable *X* is the measure of degree of dispersion, or spread, of a probability distribution about its mean (i.e. how much on average each of the values of *X* deviates from the mean.). Note: A probability distribution with a small (large) spread about its mean will have a small (large) variance.

Variance of a Random Variable X

Suppose a random variable has the probability distribution

x	x_1	x_2	…	x_n
P(X=x)	p_1	p_2	…	p_n

and expected value E(X) = μ. Then the **variance** of the random variable *X* is

$$\boxed{Var(X) = p_1(x_1-\mu)^2 + p_2(x_2-\mu)^2 + \ldots + p_n(x_n-\mu)^2}$$

Note: We square each since some may be negative.

Standard Deviation measures the same thing as the variance. The standard deviation of a random variable X is

$$\boxed{\sigma = \sqrt{Var(X)}}$$

Note: As stated above the standard deviation measures the same thing as the variance. Since we square the $(x_i-\mu)$s in the variance formula the units of *x* are squared, so to remedy this we take the square root.

Example 1: Given the following probability distribution, find its mean, variance and standard deviation.

x	P(X=x)
-2	0.25
1	0.58
3	0.17

$E(x) = 0.25(-2) + .58(1) + .17(3) = .59$

$Var(x) = .25(-2-.59)^2 + .58(1-.59)^2$

$+ .17(3-.59)^2 = 2.7619$

$\sigma = \sqrt{2.7619} = 1.6619$

Example 2: You record the length of time in minutes it takes you to drive from your home to school in 10 consecutive days. The probability distribution follows. Let the random variable X denote the number of minutes it takes you to drive from home to school. Calculate the mean and standard deviation.

x	P(X=x)
45	0.2
48	0.3
51	0.5

Mean:

$E(x) = .2(45) + .3(48) + .5(51) = 48.9$

Standard Deviation:

$Var = .2(45-48.9)^2 + .3(48-48.9)^2 + 0.5(51-48.9)^2$

$= 5.49 \qquad \sigma = \sqrt{5.49} = 2.34$

Example 3: An investor is considering two business ventures. The anticipated returns (in thousands of dollars) of each venture are described by the following probability distributions:

Venture 1

Earnings	Probability
-5	0.2
30	0.6
60	0.2

Venture 2

Earnings	Probability
-15	0.15
50	0.75
100	0.10

a. The mean of Venture 1 is 29 and the mean of Venture 2 is 45.25. Compute the standard deviation for each venture.

Venture 1:

$Var = .2(-5-29)^2 + .6(30-29)^2 + .2(60-29)^2 = 424$

$\sigma = \sqrt{424} = 20.59$

Venture 2:

$Var = .15(-15-45.25)^2 + .75(50-45.25)^2 + .1(100-45.25)^2$

$= 433.6875$

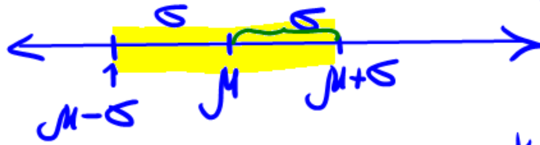
$\sqrt{433.6875} = 20.83$

b. Which investment would provide the investor with the higher expected return?

Venture 2 greater mean

c. Which investment would the element of risk be less?

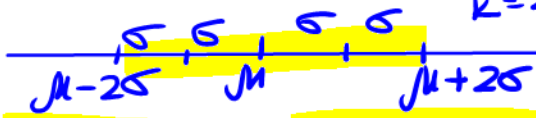
Venture 1 smaller s.d.



$k=1$

$P \geq 1 - \frac{1}{1^2} = 0$

Chebychev's Inequality



$k=2$

$P \geq 1 - \frac{1}{2^2} = \frac{3}{4}$

Let X be a random variable with expected value μ and standard deviation σ. Then, the probability that a randomly chosen outcome of the experiment lies between $\mu - k\sigma$ and $\mu + k\sigma$ is at least $1 - \frac{1}{k^2}$; that is, $P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$.

Example 4: A probability distribution has a mean 20 and a standard deviation of 3. Use Chebychev's Inequality to estimate the probability that an outcome of the experiment lies between 12 and 28.

$\mu = 20 \quad \sigma = 3$

$P(12 \leq X \leq 28) \geq 1 - \frac{1}{\left(\frac{8}{3}\right)^2} = 1 - \frac{1}{\frac{64}{9}}$

$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$

$= 1 - \frac{9}{64}$

$12 = 20 - k(3) \qquad 28 = 20 + k(3)$

$12 = 20 - 3k \qquad 28 = 20 + 3k$

$-8 = -3k \qquad 8 = 3k$

$k = 8/3 \longleftrightarrow k = 8/3$

$= \frac{64}{64} - \frac{9}{64}$

$= \boxed{\frac{55}{64}}$

Example 5: The Great Northwestern Lumber Company employs 400 workers in its mills. It has been estimated that X, the random variable measuring the number of mill workers who have industrial accidents during a 1- year period, is distributed with a mean of 30 and a standard deviation of 4. Use Chebychev's Inequality to estimate the probability that the number of workers who will have an industrial accident over a 1-year period is between 20 and 40, inclusive.

$\mu = 30 \quad \sigma = 4$

$P(\overset{\mu-k\sigma}{20} \leq X \leq \overset{\mu+k\sigma}{40}) \geq 1 - \frac{1}{\left(\frac{5}{2}\right)^2} = 1 - \frac{1}{\frac{25}{4}}$

$20 = 30 - k(4) \qquad 40 = 30 + k(4)$

$20 = 30 - 4k \qquad 40 = 30 + 4k$

$= 1 - \frac{4}{25}$

$-10 = -4k \qquad 10 = 4k$

$= \frac{25}{25} - \frac{4}{25}$

$k = \frac{10}{4} \qquad k = \frac{10}{4}$

$= \frac{21}{25} = \boxed{.84}$

$k = \frac{5}{2} \qquad k = \frac{5}{2}$