Stat 400, section 3.3 Expected Value, Variance and Standard Deviation

notes by Tim Pilachowski

Do you remember how to calculate an average? The word "average", however, has connotations outside of a strict mathematical definition, so mathematicians have a different name: the ***expected value*** or ***mean***.

Example A: Suppose we measure the heights of 25 people to the nearest inch and get the following results:

height (in.)	64	65	66	67	68	69	70
frequency	5	7	6	4	1	0	2

What is the expected value for height? *answer*: 65.88 in.

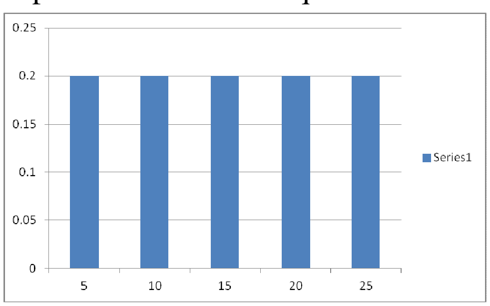
method #1:

method #2:

method #3:

$E(X) = \mu_X = x_1\,p(x_1) + x_2\,p(x_2) + x_3\,p(x_3) + \ldots + x_n\,p(x_n) = \sum_x x_i\,p(x_i)$ where x_i is a possible value for a random variable X and $p(x_i)$ is its probability.

Examples B: The four histograms below represent four sets of data. a) I'll leave it to you to show that the expected value/mean equals 15 for each of them. b) We'll come back to $V(X)$ and σ later.

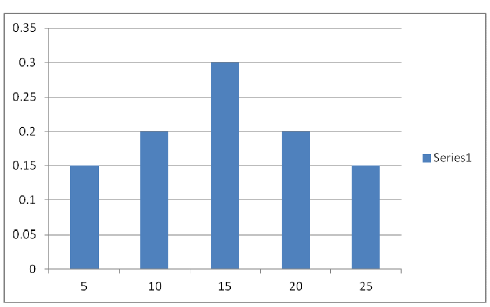


x	5	10	15	20	25
$P(X = x)$					

$E(X) = \mu_X =$

$V(X) = \sigma_X^2 =$

$\sigma_X =$

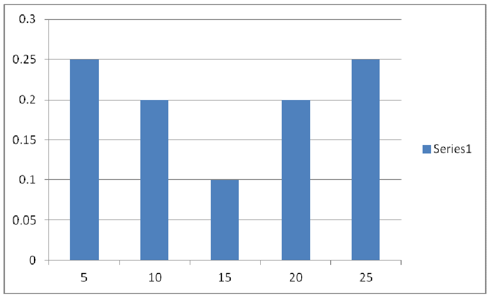


x	5	10	15	20	25
$P(X = x)$					

$E(X) = \mu_X =$

$V(X) = \sigma_X^2 =$

$\sigma_X =$

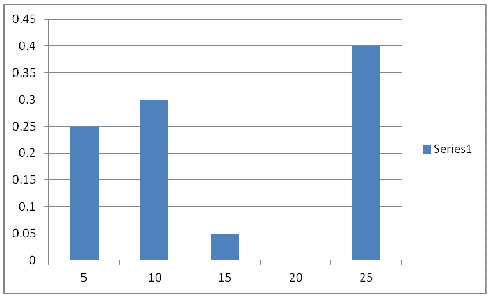


x	5	10	15	20	25
$P(X = x)$					

$E(X) = \mu_X =$

$V(X) = \sigma_X^2 =$

$\sigma_X =$



x	5	10	15	20	25
$P(X = x)$					

$E(X) = \mu_X =$

$V(X) = \sigma_X^2 =$

$\sigma_X =$

The mean/expected value alone tells about the middle, but nothing about the shape of the distribution, i.e. how the data values are distributed. Enter two new, but related concepts: *variance* and *standard deviation*.
For a probability distribution with expected value $E(X) = \mu$,

$$V(X) = \sigma_X^2 = \sigma^2 = (x_1 - \mu)^2 * p(x_1) + (x_2 - \mu)^2 * p(x_2) + \ldots + (x_n - \mu)^2 * p(x_n) = \sum_x (x_i - \mu)^2 * p(x_i)$$

Variance can be thought of as "sum of [probability times (value minus mean) squared] ".

However, since the calculations using this formula can become quite onerous, we're going to mostly use a shortcut formula for variance:

$$V(X) = \sigma^2 = \left[\sum_x (x_i)^2 * p(x_i)\right] - \mu^2 = E\left(X^2\right) - [E(X)]^2$$

Once we have $V(X)$, the standard deviation of $X = \sigma_X = \sqrt{\sigma_X^2}$.

Variance and standard deviation are both measures of how much the amounts (x_i) vary (or deviate) from the mean ($E(X) = \mu$).

Examples B revisited: b) Go back to the four sets of data in Example B and calculate the variance and standard deviation for each of them. For the first set of data we'll use both the long and the shortcut formula. For the rest, we'll use only the shortcut formula.

notes about the relationship between standard deviation and the shape of the probability distribution:

Example A revisited: We measure the heights of 25 people to the nearest inch and get the following results:

x = height (in.)	64	65	66	67	68	69	70
$p(x)$ = Probability							

What is the expected value, $E(X)$, for height? What are $V(X)$ and the standard deviation for these heights?
answers: 65.88 in., ≈ 2.6656, ≈ 1.6327

Example C: Insurance companies use actuarial data to set rates for policies. Collected data indicate that, on a \$1000 policy, an average of 1 in every 100 policy holders will file a \$20,000 claim. An average of 1 in every 200 policy holders will file a \$50,000 claim. An average of 1 in every 500 policy holders will file a \$100,000 claim. What are $E(X)$, $V(X)$ and σ for the value of a policy to the company? If the company sells 100,000 policies, what is its expected profit or loss?

x = value of a policy					
$p(x)$					Total **= 1**

answers: \$350, 36077500, ≈ 6006.4549, \$35,000,000

Example D. You test electrical components to determine whether they work or not. Define

$$X = \begin{cases} 1 & \text{if the component works (success)} \\ 0 & \text{if the component fails} \end{cases} \quad \text{with probability density function } p(x) = \begin{cases} 0.1 & \text{if } x = 0 \\ 0.9 & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Calculate $E(X) = \mu$, $E(X^2)$, $V(X) = \sigma^2$ and σ. *answers*: 0.9, 0.9, 0.09, 0.3