IX. Variance and Standard Deviation

Definition of Variance

- The variance of a random variable Y is
$$\underline{\sigma^2(Y) = V(Y)} := E\left[(Y-\mu)^2\right],$$
where $\mu := E(Y)$ is the mean (expected value) of the variable.
- Variance is a measure of the variability, or volatility, in Y.
- The most useful way to calculate variance is as follows:
$$\boxed{V(Y) = E\left(Y^2\right) - E(Y)^2}$$

Graph one smooth line and one wobbly line with the same mean but different variances.

Definition of Standard Deviation

- The standard deviation of a random variable Y is
$$\underline{\sigma} := \sqrt{V(Y)}.$$
- Because it is defined directly from variance, standard deviation also measures volatility in Y.

Example I

Which of the following sets of data has the largest variance?

A. {123, 123, 123, 123, 123}

B. {11, 11, 13, 15, 15}

C. {20, 21, 22, 23, 24}

$\boxed{\text{B}}$

Example II

Which of the following would be the least useful in understanding a set of data?

A. Knowing the mean and the standard deviation.

B. Knowing the mean and the variance.

C. Knowing the standard deviation and the variance.

$\boxed{\text{C}}$

Example III

In your probability class, the two midterm exams count for 25% each of the semester grade, the final exam counts for 30%, and the homework counts for 20%. You score 79 on both midterms, the final, and the homework. Calculate the mean, variance, and standard deviation of your scores, using the weights above.

$$\begin{aligned}
\mu &= \sum_y p(y)y = \boxed{79} \\
\sigma^2 &= E\left[(Y-\mu)^2\right] = \boxed{0} \\
\sigma &= \boxed{0}
\end{aligned}$$

Example IV

In your probability class, the two midterm exams count for 25% each of the semester grade, the final exam counts for 30%, and the homework counts for 20%. You score 60 on the first midterm, 80 on the second, 80 on the final, and 100 on the homework. Calculate the mean, variance, and standard deviation of your scores, using the weights above.

$$\mu := \sum_{y \in \mathbb{R}} p(y)y = \underbrace{60}_{y} \cdot \underbrace{\frac{25}{100}}_{p(y)} + \underbrace{80}_{y} \cdot \underbrace{\frac{55}{100}}_{p(y)} + \underbrace{100}_{y} \cdot \underbrace{\frac{20}{100}}_{p(y)} = \text{exactly } \boxed{\frac{79}{100}}$$

$$\begin{aligned}
\sigma^2 &= E\left(Y_{\text{you}}^2\right) - \mu_{\text{you}}^2 \\
&= \left(\sum y^2 p(y)\right) - \mu_{\text{you}}^2 \\
&= 60^2 \cdot \frac{25}{100} + 80^2 \cdot \frac{55}{100} + 100^2 \cdot \frac{20}{100} - 79^2 \\
&= 6420 - 79^2 \\
&= \boxed{179} \\
\sigma &= \sqrt{179} \approx \boxed{13}
\end{aligned}$$

Check:

$$E\left[(Y-\mu)^2\right] = \sum_y p(y)(y-\mu)^2 = \frac{25}{100}(19)^2 + \frac{55}{100}1^2 + \frac{20}{100}(21)^2 = \frac{17900}{100} = \boxed{179} \checkmark$$

Example V

Roll one die and let Y be the number showing. Calculate the mean, variance, and standard deviation of Y.

$$\begin{aligned}
\mu &= \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = \boxed{\frac{7}{2}} \\
E\left(Y^2\right) &= \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6} \\
\sigma^2 &= E\left(Y^2\right) - E(Y)^2 = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \boxed{\frac{35}{12}} \\
\sigma &= \boxed{\sqrt{\frac{35}{12}}}
\end{aligned}$$

Example VI

Flip a coin three times and let Y be the number of heads. Calculate the mean, variance, and standard deviation of Y.

outcome	Y	Y^2
HHH	3	9
HHT	2	4
HTH	2	4
HTT	1	1
THH	2	2
THT	1	1
TTH	1	1
TTT	0	0

$$\begin{aligned}
E(Y) &= \frac{1}{8}(3+2+2+1+2+1+1+0) = \frac{12}{8} = \boxed{\frac{3}{2}} \\
E\left(Y^2\right) &= \frac{1}{8}(9+4+4+1+4+1+1+0) = \frac{24}{8} = 3 \\
\sigma^2 &= E\left(Y^2\right) - E(Y)^2 = 3 - \frac{9}{4} = \boxed{\frac{3}{4}} \\
\sigma &= \boxed{\frac{\sqrt{3}}{2}}
\end{aligned}$$