Properties of Expectation, Variance, and Standard Deviation

↳ applies to discrete and continuous random variables

Shifting:

add (or subtract) a constant to every value of a random variable X.

$E(X+c) = E(X)+c$

$E(X-c) = E(X)-c$

Let c be a constant.

- $E(X \pm c) = E(X) \pm c$ Centre
- $\text{Var}(X \pm c) = \text{Var}(X)$
- $SD(X \pm c) = SD(X)$

$\left.\begin{array}{l}\text{Var}(X \pm c) = \text{Var}(X)\\ SD(X \pm c) = SD(X)\end{array}\right\}$ spread

Example: Card Game: add \$2 to all payouts $Y = X + 2$

Outcome	y ($x+2$)	$P(y)$
Black	2	$\frac{1}{2}$
Diamond	7	$\frac{1}{4}$
Heart (not queen)	12	$\frac{3}{13}$
Queen of Hearts	32	$\frac{1}{52}$

$E(X) = \$4.13$

$\text{Var}(X) \simeq 29.54$

$SD(X) \simeq \$5.44$

$E(X+2) = E(X) + 2$
$= \$4.13 + 2 = \$6.13$

$\text{Var}(X+2) = \text{Var}(X)$
$\simeq 29.54$

$SD(X+2) = SD(X)$
$\simeq \$5.44$

Scaling: multiply every value of a random variable X by a constant.

Let a be a constant.

- $E(aX) = aE(X)$ Centre
- $\text{Var}(aX) = a^2\,\text{Var}(X)$
- $SD(aX) = |a|SD(X)$

spread

absolute value

For $a \geq 0$,

$SD(aX) = a\,SD(X)$

Example: Card Game: multiply all payouts by 4 $Y = 4X$

Outcome	y (4x)	$P(y)$
Black	0	$\frac{1}{2}$
Diamond	20	$\frac{1}{4}$
Heart (not queen)	40	$\frac{3}{13}$
Queen of Hearts	120	$\frac{1}{52}$

$E(X) \simeq \$4.13$

$\text{Var}(X) \simeq 29.54$

$SD(X) \simeq \$5.44$

$E(4X) = 4E(X)$

$= 4(4.13) = \$16.52$

$\text{Var}(4X) = 4^2\,\text{Var}(X)$

$= 16(29.54)$

$= 472.64$

$SD(4X) = 4\,SD(X)$

$= 4(5.44)$

$\simeq \$21.76$

Adding / Subtracting Random Variables

Let X be a random variable with mean / expected value $\mu_X = E(X)$, variance $\sigma_X^2 = \text{Var}(X)$, and standard deviation $\sigma_X = SD(X)$.

Let Y be a random variable with mean / expected value $\mu_Y = E(Y)$, variance $\sigma_Y^2 = \text{Var}(Y)$, and standard deviation $\sigma_Y = SD(Y)$.

Then $X+Y$ and $X-Y$ are random variables with means (expected values)

$$E(X+Y) = E(X) + E(Y) \qquad \text{and} \qquad E(X-Y) = E(X) - E(Y)$$

μ_{X+Y} $\qquad$ μ_{X-Y}

If X and Y are **independent**, then the variances are

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) \qquad \text{and} \qquad \text{Var}(X-Y) = \text{Var}(X) + \text{Var}(Y)$$

σ^2_{X+Y}

outcome of one variable tells us nothing about the likely outcome of the other variable.

Warning: In general,

$\sqrt{a+b} \neq \sqrt{a} + \sqrt{b}$

$$SD(X+Y) \neq SD(X) + SD(Y)$$

$$SD(X+Y) = \sqrt{\text{Var}(X+Y)} = \sqrt{\text{Var}(X) + \text{Var}(Y)} = \sqrt{(SD(X))^2 + (SD(Y))^2}$$

Let a, b, and c be constants and let X and Y be random variables (discrete or continuous), then

$$E(aX + bY + c) = aE(X) + bE(Y) + c$$

and if X and Y are independent

$$\text{Var}(aX + bY + c) = a^2\,\text{Var}(X) + b^2\,\text{Var}(Y)$$

Warning: For repeated instances of a random variable X, we use the notation $X_1, X_2, X_3, \ldots$, since each instance represents a different outcome of X. In particular,

$$X_1 + X_2 + X_3 \neq 3X$$

Note:

- The probability model for the sum of two random variables is not necessarily the same as the one we started with, even if the random variables are independent.
- For a continuous random variable X with a **Normal** probability model, shifting or scaling X produces a random variable that also has a Normal probability model, but with a different mean and standard deviation.
- If X and Y are **independent** continuous random variables with **Normal** probability models, then their sum $X + Y$ (or difference $X - Y$) also has a Normal probability model (with a new mean and standard deviation).

Example: Flip a fair coin.

X = number of heads observed. $E(X) = \frac{1}{2}$ $\quad \text{Var}(X) = \frac{1}{4}$

Outcome	x	$P(x)$
T	0	$\frac{1}{2}$
H	1	$\frac{1}{2}$

0

1

$SD(X) = \sqrt{Var(X)}$
$= \frac{1}{2}$

Flip two fair coins: consider the random variable $X_1 + X_2$.

independent

Outcome	$x_1 + x_2$	$P(x_1 + x_2)$
TT	0	$\frac{1}{4}$
HT or TH	1	$\frac{1}{2}$
HH	2	$\frac{1}{4}$

$E(X_1 + X_2) = E(X_1) + E(X_2)$
$= \frac{1}{2} + \frac{1}{2} = 1$

$Var(X_1 + X_2) = Var(X_1) + Var(X_2)$
$= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$SD(X_1 + X_2) = \sqrt{Var(X_1 + X_2)} = \frac{1}{\sqrt{2}}$

$\neq$

$SD(X_1) + SD(X_2) = \frac{1}{2} + \frac{1}{2} = 1$

Example: Suppose you are given two independent random variables, X and Y, with means and standard deviations:

	Mean	Standard deviation	Variance
X	10	3	9
Y	25	5	25

Compute:

a) $E(2X-1) = E(2X) - 1 = 2E(X) - 1$
$= 2(10) - 1$
$= 19$

b) $\text{Var}(2X-1) = \text{Var}(2X) = 2^2\,\text{Var}(X)$
$= 4(9)$
$= 36$

c) $E(4X-3Y+6) = E(4X-3Y)+6$
$= E(4X) - E(3Y) + 6$
$= 4E(X) - 3E(Y) + 6$
$= 4(10) - 3(25) + 6$
$= -29$

d) $\text{Var}(4X-3Y+6) = \text{Var}(4X-3Y)$
$= \text{Var}(4X) + \text{Var}(3Y)$
$= 4^2\,\text{Var}(X) + 3^2\,\text{Var}(Y)$
$= 16(9) + 9(25)$
$= 369$

Example: A company sells flower seeds in packets of 20. They estimate that the mean number of good seeds (that is, seeds that actually grow) in each packet is 18 with a standard deviation of 1.2. You buy 5 packets of seeds. X = # good seeds in a packet

X_1, X_2, X_3, X_4, X_5

$E(X) = 18 \qquad SD(X) = 1.2$

a) What is your expected number of good seeds?

$$E(X_1+X_2+X_3+X_4+X_5) = E(X_1)+E(X_2)+E(X_3)+E(X_4)+E(X_5)$$
$$= 18+18+18+18+18$$
$$= 90$$

$\sqrt{ab} = \sqrt{a}\sqrt{b}$

b) What is the standard deviation of the number of good seeds?

$$SD(X_1+X_2+X_3+X_4+X_5) = \sqrt{Var(X_1+X_2+X_3+X_4+X_5)}$$
$$= \sqrt{Var(X_1)+Var(X_2)+Var(X_3)+Var(X_4)+Var(X_5)}$$
$$= \sqrt{(1.2)^2+(1.2)^2+(1.2)^2+(1.2)^2+(1.2)^2}$$
$$= \sqrt{5(1.2)^2} = \sqrt{5}(1.2) \approx 2.683$$

$\cancel{SD(5X)}$
$= 5\,SD(X)$
$= 5(1.2)$
$= 6$

c) What assumption are you making about the seeds to answer part b)? Is this assumption reasonable?

We are assuming that the trials of picking seed packets are independent.

If the seed packets are for different types, the assumption is reasonable.

If they are all one type, they may have come from the same batch → possible dependence.

d) In one packet of seeds, what is the expected number of bad seeds? The standard deviation of the number of bad seeds?

Y = # bad seeds in a packet $= 20 - X$

$$E(Y) = E(20-X) = 20 - E(X) = 20 - 18 = 2$$

$$SD(Y) = \sqrt{Var(20-X)} = \sqrt{Var(-X)} = \sqrt{Var(X)}$$
$$= \sqrt{(1.2)^2}$$
$$= 1.2$$

Example: A farming couple send their son Bob every Saturday to a local farmer's market to sell carrots and broccoli, where he sells an average of \$250 worth of carrots with a standard deviation of \$30 and an average of \$110 worth of broccoli with a standard deviation of \$10. Suppose Bob's parents pay him \$100 plus 10% commission on the carrot sales and 15% commission on the broccoli sales for that day.

X = \$ amount of carrots sold

Y = \$ amount of broccoli sold

$E(X) = \$250$, $SD(X) = \$30$ $\qquad$ $E(Y) = \$110$, $SD(Y) = \$10$

assume independence

$I = 100 + 0.1X + 0.15Y$

a) Find the mean of Bob's Saturday income.

$$
\begin{aligned}
E(I) &= E(100 + 0.1X + 0.15Y) \\
&= 100 + 0.1E(X) + 0.15E(Y) \\
&= 100 + 0.1(250) + 0.15(110) \\
&= 100 + 25 + 16.5 = \$141.50
\end{aligned}
$$

b) Find the variance and standard deviation of Bob's Saturday income.

$$
\begin{aligned}
Var(I) &= Var(100 + 0.1X + 0.15Y) \\
&= (0.1)^2 Var(X) + (0.15)^2 Var(Y) \\
&= 0.01(30)^2 + 0.0225(10)^2 \\
&= 9 + 2.25 = 11.25
\end{aligned}
$$

$$SD(I) = \sqrt{11.25} \simeq \$3.35$$

c) What would be considered an unusually good Saturday for Bob?
A great Saturday?

good: $E(I) + \underline{2}\,SD(I) = \$141.50 + 2(\$3.35) = \$148.20$ (or more)

great: $E(I) + \underline{3}\,SD(I) = \$141.50 + 3(\$3.35) = \$151.55$ (or more)

Example: Suppose that the heights of adult males in Canada can be described by a Normal model with a mean of 176 cm and standard deviation of 7 cm and the heights of adult females in Canada can be described by a Normal model with a mean of 163 cm and standard deviation of 6 cm. Suppose we randomly select a Canadian male and, independently, a Canadian female. What is the probability that the male is taller than the female?

X = height of female

$E(X) = 163$, $SD(X) = 6$

$X \sim N(163, 6)$

Y = height of male

$E(Y) = 176$, $SD(Y) = 7$

$Y \sim N(176, 7)$

$E(Y-X) = E(Y) - E(X) = 176 - 163 = 13$

$SD(Y-X) = \sqrt{Var(Y-X)} = \sqrt{Var(Y) + Var(X)}$

assume independence $= \sqrt{7^2 + 6^2} = \sqrt{85}$

$\simeq 9.22$

$\therefore Y - X \sim N(13, 9.22)$

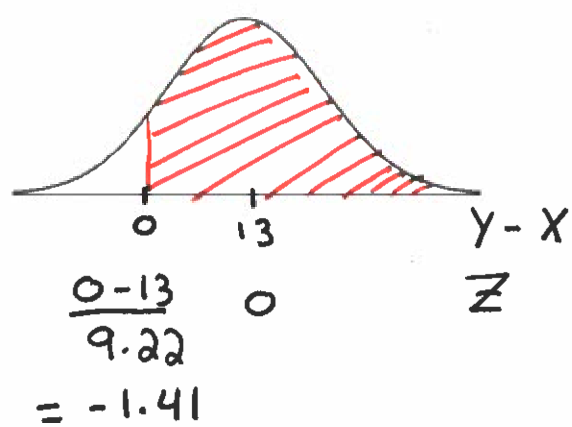
$P(Y - X > 0)$

$= P\left(Z > \frac{0-13}{9.22}\right)$

$= P(Z > -1.41)$

$= P(Z < 1.41)$

$= 0.9207$



Formulas:

Let X and Y be random variables (continuous or discrete).

Expected Value

- $E(X + c) = E(X) + c$
- $E(aX) = aE(X)$
- $E(X + Y) = E(X) + E(Y)$
- $E(aX + bY + c) = aE(X) + bE(Y) + c$

Variance

- $\mathrm{Var}(X + c) = \mathrm{Var}(X)$
- $\mathrm{Var}(aX) = a^2\,\mathrm{Var}(X)$

If X and Y are independent

- $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$
- $\mathrm{Var}(aX + bY + c) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y)$

Standard Deviation

- $SD(X + c) = SD(X)$
- $SD(aX) = |a|SD(X)$

If X and Y are independent

- $SD(X + Y) = \sqrt{\left(SD(X)\right)^2 + \left(SD(Y)\right)^2}$