

AP Statistics - Random Variables, Expected Value and Combining Variables - Lesson

Common Core AP STAT 4.7, AP STAT 4.8, AP STAT 4.9, HSS-MD.A.1, HSS-MD.A.2, HSS-MD.B.5 · maths-pdf.fr/en/

A discrete random variable has numerical values with probabilities summing to 1. Its mean, the expected value, is the sum of each value times its probability. For $Y = a + bX$ the mean is $a + b$ times the old mean and the standard deviation is $|b|$ times the old one. For independent variables, means add and variances add.

AP Statistics · Common Core State Standards: **AP STAT 4.7, AP STAT 4.8, AP STAT 4.9, HSS-MD.A.1, HSS-MD.A.2, HSS-MD.B.5**

Key facts

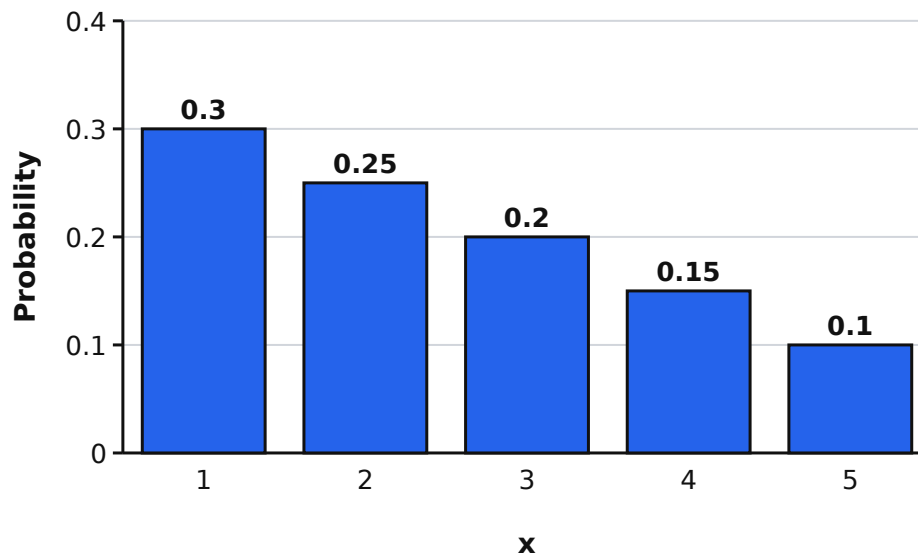
- Probabilities in a distribution are between 0 and 1 and sum to 1.
- Mean (expected value): sum of $x p_i$; interpret it as the long-run average.
- Standard deviation: square root of the sum of $(x - \text{mean})^2 p_i$.
- For $a + bX$: mean $a + b \mu$, standard deviation $|b| \sigma$.
- For independent X, Y : means add or subtract; variances add (also for $X - Y$).

What is a discrete random variable?

A **random variable** X is a numerical outcome of a random process. A **discrete** random variable has a countable list of possible values, and its **probability distribution** gives the probability of each value. The probabilities must be between 0 and 1 and must add to 1.

x	1	2	3	4	5
P(x)	0.3	0.25	0.2	0.15	0.1

Probability distribution of X , the number of items in an order.



Probability histogram: skewed right, with mean 2.5 items.

Reading a probability histogram works like reading any histogram. The tallest bar marks the most likely value, here 1 item, and the long tail on the right shows that big orders are possible but rare. The mean, 2.5 items, lies to the right of the tallest bar because of this skew, just as the mean of a skewed data set lies toward the tail. Always check the shape before you decide how to describe the distribution in context.

How do you find the mean and standard deviation of a random variable?

The **mean** (expected value) is the long-run average: $\mu_X = \sum x_i p_i$. The **standard deviation** is $\sigma_X = \sqrt{\sum (x_i - \mu_X)^2 p_i}$.

x	p	x p	(x - 2.5) ² p
1	0.3	0.30	0.6750
2	0.25	0.50	0.0625
3	0.2	0.60	0.0500
4	0.15	0.60	0.3375
5	0.1	0.50	0.6250
Sum	1	2.50	1.7500

Mean 2.50; variance 1.7500; standard deviation 1.323.

For the order data, $\mu_X = 2.50$ items and $\sigma_X = 1.323$ items. The mean is not a value that must occur; it tells us that over many orders the average is 2.5 items per order.

How do transformations and combinations change a random variable?

Operation	Mean	Standard deviation
$a + bX$	$a + b \mu$	$ b \sigma$
$X + Y$	$\mu_X + \mu_Y$	$\sqrt{s_x^2 + s_y^2}$
$X - Y$	$\mu_X - \mu_Y$	$\sqrt{s_x^2 + s_y^2}$

Rules for transforming and combining random variables (X and Y independent).

Suppose shipping costs $Y = 3 + 2X$ dollars. Then $\mu_Y = 3 + 2(2.5) = 8.0$ and $\sigma_Y = 2(1.323) = 2.646$ dollars. Adding 3 does not change the spread.

Combining independent variables

Means always add or subtract. Variances add for both sums and differences when the variables are independent: the standard deviation of $X - Y$ is the same as for $X + Y$.

If X and Y are independent with $\sigma_X = 3$ and $\sigma_Y = 4$, then $\sigma_{X+Y} = \sqrt{3^2 + 4^2} = 5$, not 7. Standard deviations do not add; variances do.

1. Write the distribution and check the probabilities add to 1.
2. Compute the mean, then the variance and standard deviation.
3. Apply the transformation or combination rule.
4. Interpret the mean and standard deviation in context.

Questions about random variables, expected value and combining variables

What does the expected value mean?

It is the long-run average value of the random variable over many repetitions. It need not be a possible value: an average of 2.5 items per order is perfectly sensible even though no order has 2.5 items.

Why do we add variances instead of standard deviations?

For independent variables the variability combines like the sides of a right triangle, so variances add. Standard deviations add only when the variables are perfectly related.

Is the standard deviation of $X - Y$ smaller than that of $X + Y$?

No. For independent variables, sigma of $X - Y$ equals sigma of $X + Y$, since the variances add in both cases. Subtracting does not reduce the variability.

What is the difference between $2X$ and $X + X$?

In $2X$, one value of X is doubled, so the standard deviation is 2σ . If two independent copies X and X are added, the standard deviation is $\sqrt{2}\sigma$, which is smaller.

How do I decide whether a game is fair?

A game is fair if the expected net gain is 0. Compute the expected value of the net winnings, including the cost to play.