

What is the definition and an example of the 10% condition in statistics?

Authored by
stats writer

April 19, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the definition and an example of the 10% condition in statistics?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=137095>

The 10% condition in statistics refers to a guideline that states that the sample size used in a statistical analysis should not exceed 10% of the total population size. This condition is important in order to ensure that the sample is truly representative of the population and that the results of the analysis can be generalized to the entire population. For example, if a researcher wants to study the average height of all high school students in a particular city, the sample size should not be more than 10% of the total number of high school students in that city. This ensures that the sample is not biased and accurately reflects the characteristics of the entire population.

The 10% Condition in Statistics: Definition & Example

A Bernoulli trial is an experiment with only two possible outcomes - "success" or "failure" - and the probability of success is the same each time the experiment is conducted.

An example of a Bernoulli trial is a coin flip. The coin can only land on two sides (we could call heads a "success" and tails a "failure") and the probability of success on each flip is 0.5, assuming the coin is fair.

Often in statistics when we want to calculate probabilities involving more than just a few Bernoulli trials, we use the normal distribution as an approximation. However, in order to do so we must assume that the trials are independent.

In cases where the trials are *not* actually independent, we can still assume that they are if the sample size

we're working with does not exceed 10% of the population size. This is known as The 10% Condition.

The 10% Condition: As long as the sample size is less than or equal to 10% of the population size, we can still make the assumption that Bernoulli trials are independent.

Intuition Behind The 10% Condition

To develop an intuition behind The 10% Condition, consider the following example.

Suppose the true proportion of students in a certain class who prefer football over basketball is 50%. Let random variable X be the number of students randomly selected in 4 trials who prefer football over basketball. Let's say we're interested in understanding the probability that all 4 randomly selected students prefer football over basketball.

If our classroom size is 20 and our trials were independent (e.g. we could take repeated samples of all 20 students), then the probability that each student would prefer football over basketball could be calculated as:

$P(\text{All 4 students prefer football}) = 10/20 * 10/20 * 10/20 * 10/20 = .0625.$

However, if our trials are *not* independent (e.g. once we sample one student, they can't be placed back in the classroom) then the probability that all 4 students would prefer football would be calculated as:

$P(\text{All 4 students prefer football}) = 10/20 * 9/19 * 8/18 * 7/17 = .0433.$

These two probabilities are quite different. Consider that in this example our sample size (4 students) is not less than or equal to 10% of the population (20 students), thus we wouldn't be able to use The 10% Condition.

However, consider the following table that shows the probability that all 4 randomly selected students prefer football, based on classroom size:

Classroom Size	Sample Size / Classroom Size	P(All 4 students prefer football) <i>Independent trials</i>	P(All 4 students prefer football) <i>Not Independent trials</i>
20	20.0%	$10/20 * 10/20 * 10/20 * 10/20 = .0625$	$10/20 * 9/19 * 8/18 * 7/17 = .0433$
40	10.0%	$20/40 * 20/40 * 20/40 * 20/40 = .0625$	$20/40 * 19/39 * 18/38 * 17/37 = .0530$
100	4.0%	$50/100 * 50/100 * 50/100 * 50/100 = .0625$	$50/100 * 49/99 * 48/98 * 47/97 = .0587$
1,000	0.4%	$500/1,000 * 500/1,000 * 500/1,000 * 500/1,000 = .0625$	$500/1,000 * 499/999 * 498/998 * 497/997 = .0621$

As the sample size relative to the population size (e.g. "classroom size" in this example) decreases, the calculated probability between independent trials and non-independent trials gets closer and closer.

Note that when the sample size is exactly 10% of the population size, the difference between the probabilities of independent trials and non-independent trials are relatively similar.

Conclusion

The 10% Condition says that our sample size should be less than or equal to 10% of the population size in order to safely make the assumption that a set of Bernoulli trials is independent.

Of course, it's best if our sample size is much less than 10% of the population size so that our inferences about the population are as accurate as possible. For example, we'd prefer that our sample size is only 5% of the population compared to 10%.

[An Introduction to the Normal Distribution](#)

[An Introduction to the Binomial Distribution](#)

An Introduction to the Central Limit Theorem

ARABPSYCHOLOGY.COM