Normal Approximation to the Binomial

Can be done when: 1. N is large, 2. Probability of success is not too close to 0 or 1

The normal approximation to the binomial is when you use a continuous distribution (the normal distribution) to approximate a discrete distribution (the binomial distribution). According to the **Central Limit Theorem**, the sampling distribution of the sample means becomes approximately normal if the sample size is large enough.

The Normal Approximation to the Binomial Distribution is used when certain conditions are met, primarily when the number of trials (n) in the binomial distribution is large, and the probability of success (p) is not too close to 0 or 1.

The Normal Distribution is a continuous probability distribution that is well understood and has various properties that make it easier to work with in many cases. When the number of trials is large, the shape of the Binomial Distribution approaches that of a bell-shaped curve, which is a characteristic of the Normal Distribution. This allows us to use the Normal Distribution to estimate probabilities and calculate confidence intervals for the Binomial Distribution.

The first step into using the normal approximation to the binomial is making sure you have a “large enough sample”. How large is “large enough”? You figure this out with two calculations: n \* p and n \* q .

n is your sample size,

p is your given probability.

q is just 1 – p.

When **n \* p and n \* q, are greater than 5**, you can use the normal approximation to the binomial to solve a problem.

If X is a random variable that follows a binomial distribution with n trials and p probability of success on a given trial, then we can calculate the mean (μ) and standard deviation (σ) of X using the following formulas:

Mean-mu🡪 μ = np

Standard Deviation- sigma🡪 σ = √np(1-p)

**Example: Normal Approximation to the Binomial**

Suppose we want to know the probability that a coin lands on heads less than or equal to 43 times during 100 flips.

In this situation we have the following values:

n (number of trials) = 100

X (number of successes) = 43

p (probability of success on a given trial) = 0.50

To calculate the probability of the coin landing on heads less than or equal to 43 times, we can use the following steps:

**Step 1: Verify that the sample size is large enough to use the normal approximation.**

First, we must verify that the following criteria are met:

np ≥ 5

n(1-p) ≥ 5

In this case, we have:

np = 100\*0.5 = 50

n(1-p) = 100\*(1 – 0.5) = 100\*0.5 = 50

Both numbers are greater than 5, so we’re safe to use the normal approximation.

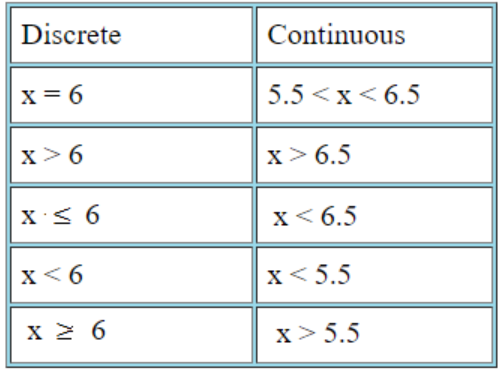
**Step 2: Determine the continuity correction to apply.**

Referring to the table below, we see that we should add 0.5 when we’re working with a probability in the form of X ≤ 43. Thus, we will be finding P(X< 43.5).

|  |  |
| --- | --- |
| Using Binomial Distribution | Using Normal Distribution with Continuity Correction |
| X = 45 | 44.5 < X < 45.5 |
| X ≤ 45 | X < 45.5 |
| X < 45 | X < 44.5 |
| X ≥ 45 | X > 44.5 |
| X > 45 | X > 45.5 |

What is Continuity Correction: <https://www.statisticshowto.com/what-is-the-continuity-correction-factor/>

Continuity corrections are adjustments that can be made to statistical tests when discontinuous distributions (like binomial distributions) are approximated by continuous ones (like normal distributions).



**Step 3: Find the mean (μ) and standard deviation (σ) of the binomial distribution.**

μ = n\*p = 100\*0.5 = 50

σ = √n\*p\*(1-p) = √100\*.5\*(1-.5) = √25 = 5

**Step 4: Find the z-score using the mean and standard deviation found in the previous step.**

z = (x – μ) / σ = (43.5 – 50) / 5 = -6.5 / 5 = -1.3

**Step 5: Find the probability associated with the z-score.**

We can use the Normal CDF Calculator to find that the area under the standard normal curve to the left of -1.3 is .0968.

Thus, the probability that a coin lands on heads less than or equal to 43 times during 100 flips is .0968.

This example illustrated the following:

* We had a situation where a random variable followed a binomial distribution.
* We wanted to find the probability of obtaining a certain value for this random variable.
* Since the sample size (n = 100 trials) was sufficiently large, we were able to use the normal distribution to approximate the binomial distribution.
* This is a complete example of how to use the normal approximation to find probabilities related to the binomial distribution.