

## **z-Test Approximation of the Binomial Test**

A binary random variable (e.g., a coin flip), can take one of two values. If we arbitrarily define one of those values as a success (e.g., heads=success), then the following formula will tell us the probability of getting  $k$  successes from  $n$  observations of the random variable when the probability of a success equals  $p$ .

$$P(k | n, p) = \binom{n}{k} p^k (1 - p)^{n-k} \quad (1)$$

The first part of the formula can be read as “ $n$  choose  $k$ ”, and is computed as follows:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (2)$$

where ! means factorial.

For example, we can use the formula to compute the probability of getting exactly two heads in three flips of a fair coin.

$n=3$  (i.e., there are three flips)

$k=2$  (i.e., heads comes up twice)

$p=.5$  (i.e., the probability of the coin landing heads up is 50%)

$$P(k = 2 | n = 3, p = .5) = \binom{3}{2} (.5)^2 (1 - .5)^{3-2} = \frac{3!}{2!(1!)} (.5)^2 (.5)^1 = \frac{6}{2(1)} (.5)^3 = 3(.125) = .375$$

However, when  $n$  becomes large enough the factorials in Equation 2 become too large to compute (e.g., my computer can't handle  $n > 69$ ). Fortunately, thanks to the central limit theorem, for such large values<sup>1</sup> of  $n$  we can accurately approximate the binomial distribution defined by Equation 1 with a normal distribution with the following mean and standard deviation:

$$\mu = np, \quad \sigma = \sqrt{np(1-p)}$$

This enables us to approximate binomial tests for a large number of observations with  $z$ -tests. Consider, for example, the following problem:

---

<sup>1</sup> When the null hypothesis is  $p=.5$  and the alpha level is .05, then  $n$  can be as small as 27. When the distribution of the null hypothesis is skewed (e.g.,  $p=.95$ ), larger sample sizes will be necessary for the normal distribution to be an accurate approximation (Zar, 1999, pgs. 535-538)

Nixon thinks that he's a better at the game "rock, paper, scissors" than his friend Kissinger. To find out if this is the case, he challenges Kissinger to 49 bouts of rock, paper, scissors. Nixon wins 31 of these bouts. Can we reject the null hypothesis that the two men are equally good at the game (i.e.,  $P(\text{Nixon wins})=P(\text{Nixon loses})=.5$ ) at an alpha level of .05?

Because our measurements are binary (Nixon wins or loses), the null hypothesis is binomially distributed with the following parameters:  $n=49$ ,  $p=.5$ . Because  $n$  is large we can approximate the distribution with a normal distribution with a mean of 24.5 and standard deviation of 3.5. We can now conduct a  $z$ -test.

Null Hypothesis	$N(\mu=24.5, \sigma=3.5)$
Alternative Hypothesis	$\mu > 24.5$
Tail of Test	upper tailed
Type of Test	$z$ -test
Alpha level	$\alpha=.05$
Critical Value(s) of Test Statistic	$z=1.65$
Observed Value of Test Statistic	$z(n=49)=1.86$
$p$ -value of Observed Value of Test Statistic	$p=.0314$
Conclusion	Reject the Null Hypothesis

1.65 is our critical  $z$ -score because just less than 5% of the area under the standard normal distribution lies between it and positive infinity (4.95% of it to be more exact). To get out test statistic, we convert the number of Nixon's victories, 31, into a  $z$ -score:

$$z = \frac{x - \mu}{\sigma} = \frac{31 - 24.5}{3.5} \approx 1.86$$

Since the  $z$ -score of our sample exceeds our critical  $z$ -score, we reject the null hypothesis that Nixon and Kissinger are equally good in favor of the alternative hypothesis that Nixon is better. To get a sense of how significant 31 victories are, we compute the  $p$ -value of our sample, 3.14%, which is the percentage of the area under the standard

normal distribution that lies between 1.86 and positive infinity. Our hypothesis test is thus concluded.

Note, another way you could have performed the binomial test is to have used the MEAN number of wins rather than the TOTAL number of wins. This might be an easier way to communicate your results (i.e., It might be easier to understand that Nixon won 63% of the time than it is to understand that he won 31 out of 49 times). If you take this approach, the mean and standard deviation of the null hypothesis are:

$$\mu = p, \sigma = \sqrt{p(1-p)}$$

and the mean and standard error of the mean (i.e., the standard deviation of the sampling distribution of the mean) are:

$$\mu_{\bar{X}} = p, \sigma_{\bar{X}} = \sqrt{\frac{p(1-p)}{n}}$$

Our observed z-score remains the same:

$$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} \approx \frac{.6327 - .5}{\frac{.5}{\sqrt{49}}} \approx \frac{.1327}{.0714} \approx 1.86$$

as does the rest of our hypothesis test:

Null Hypothesis	N( $\mu=.5, \sigma=.5$ )
Alternative Hypothesis	$\mu>.5$
Tail of Test	upper tailed
Type of Test	z-test
Alpha level	$\alpha=.05$
Critical Value(s) of Test Statistic	$z=1.65$
Observed Value of Test Statistic	$z(n=49)=1.86$
p-value of Observed Value of Test Statistic	$p=.0314$

Conclusion	Reject the Null Hypothesis
------------	----------------------------

### References

Zar, J. H. (1999). *Biostatistical Analysis* (Fourth Ed.). Upper Saddle River, New Jersey: Prentice Hall.