

Hypothesis testing; also, normal approximation to binomial

0.1 Two examples of significance testing

(I.) A certain remedy is effective in relieving an allergy for 8 hours, with probability p .

A trial of 200 people found that it did so for 160 participants. Evaluate the claim that $p = 90\%$, at the 1% level of significance.

Answer. $H_0: p = .9$; $H_1: p < .9$. This time the null hypothesis is what is claimed.

Also, we use $\text{Prob}_1, \text{Prob}_2, \text{Prob}_3$ for $B(200, 0.9)$ probability, $N(180, 18)$, probability, and $N(0, 1)$ probability distributions.

Now, we are interested in a 1-tailed low critical region of 1%. Since $\text{Prob}_3 \sim N(0, 1)$, we refer to the tables for the 1% low critical region, and get

$$(-\infty, -2.33]$$

Conversion of $N(180, 18)$ to $N(0, 1)$ by translation and scaling leads us to this critical region for Prob_2 :

$$(-\infty, 180 - 2.33\sqrt{18}] = (-\infty, 170.1146).$$

Allowing for the continuity correction, the critical region for $\text{Prob}_1 : B(200, .9)$ is as follows:

$$\text{Prob}_1(X \leq k) = \text{Prob}_2(X \leq k + 0.5) \approx 0.01$$

Whence

$$X \leq 170.1146 - 0.5 = 169.6146$$

And since X is an integer, $X \leq 169$. This is the 1% critical region. The actual outcome, 160, is well inside that region and therefore the null hypothesis is rejected.

(II.) 100 iid $N(\mu, \sigma^2)$ variables have sample mean $\bar{X} = 10.2165$ and sample standard deviation $S = 2.0114$. For a sample this large, the difference between Student's t-distribution and $N(0, 1)$, unnecessary, and it may be assumed that $\sigma = 2.0114$.

It is believed that $\mu = 10$. Make this the null hypothesis, and let the alternative hypothesis be that $\mu \neq 10$. Test this at the 1% significance level.

Answer. $H_0 : \mu = 10$. $H_1 : \mu \neq 10$.

Under the null hypothesis, $\bar{X} \sim N(\mu, \sigma^2/100)$. Choose α so if $Z \sim N(0, 1)$,

$$\text{Prob}(-\alpha \leq Z \leq \alpha) = 99 :$$

from the tables, $\alpha = 2.575$.

Then

$$\frac{\bar{X} - \mu}{\sigma/10} \in [-2.575, 2.575]$$

with 99% probability.

$\sigma/10 = .20114$, and under the null hypothesis,

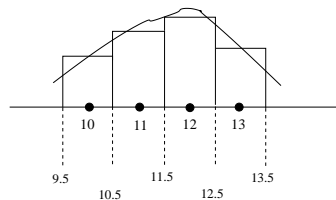
$$\bar{X} \in [10 \mp 2.575 \times .20115] = [9.482, 10.518]$$

So the result, $\bar{X} = 10.2165$, is within the noncritical region. Accept H_0 .

0.2 Continuity correction

Continuity correction improves the accuracy when a ‘discrete’ distribution is approximated by a continuous distribution. We are interested in $B(n, p)$ and $N(0, 1)$, whose sample spaces are $\{0, \dots, n\}$ and \mathbb{R} respectively.

Represent the discrete distribution by a bar diagram, as shown.



The continuity correction is based on the idea that the probability that $X = k$ (in $B(n, p)$) is the area of the rectangle based around k and should be approximated by the corresponding area under the normal curve.

With this in mind,

$$\text{Prob}_1(X \leq k) \approx \text{Prob}_2(X \leq k + 0.5)$$

and

$$\text{Prob}_1(X \geq k) \approx \text{Prob}_2(X \geq k - 0.5)$$

etcetera.

0.3 Normal approximation to the binomial

```
EXAMPLE: n,p,significance: 200 0.9 0.01
0.010000 significance N(0,1) alpha 2.330000
n 200 p 0.9000
one-tailed critical regions estimated for B(200,0.9)
```

```
low critical region 1% significance, continuity corr: [0,169.615]
```

```
Round down.
```

```
high critical region at 1% significance: [190.385,200]
```

```
Round up.
```

```
two-tailed critical region at 1% significance
```

```
Union of [0,168.575] and [191.425,200]
```

```
[0,168] union [192,200]
```

0.4 Disagreement; go with these figures

These figures were checked with some online help, and the upper 1% tail, $[191, 200]$, disagrees with the online help which would say $[190, 200]$. An experiment with $B(200, 0.1)$ reveals a disagreement but at the lower tail.

We believe that the disagreement arises from inaccuracies in the normal approximation. Anyway the difference is slight. One should continue to use the normal distribution with a continuity correction.

0.5 More examples

EXAMPLE: n,p,significance 30 0.5 0.05

0.050000 significance N(0,1) alpha 1.645000

n 30 p 0.5000

one-tailed critical regions estimated for B(30,0.5)

low critical region 5% significance, continuity corr: [0,9.99498]

high critical region at 5% significance: [20.005,30]

two-tailed critical region at 5% significance

Union of [0,9.13232] and [20.8677,30]