

# Significance Tests

## An overview

- We may view sample data as outcomes of a data generating process.
- The data generating process may be described by a probability distribution with certain parameters.
- We begin by considering a random sample  $Y_1, Y_2, \dots, Y_n$  of size  $n$  from normal distribution  $N(\mu, \sigma^2)$  with unknown  $\mu$  and known  $\sigma^2$ .
- We then use statistics from the sample data to draw inference about the parameters of the probability distribution.
- For example, use the sample mean  $\bar{Y}$  to draw inference about the population mean  $\mu$ .
- There are several ways of statistical inference. One of them is significance testing.
- The main idea of significance tests is to make a yes/no decision about some aspect of a population, based on a single sample.

## Significance Tests with Known $\sigma$

### Fruit can example

A fruit canning company claims that the average weight of a can of peaches is 15 oz, but there is some suspicion that the canner is not filling the cans. Suppose the weight of cans have a normal distribution  $N(\mu, (0.3)^2)$ . An investigator plans to take a random sample of 60 cans, weigh them, and compute the sample mean  $\bar{Y}$ . Her goal is to determine if what is observed about  $\bar{Y}$  supports the claim that the average weight of a can of peaches is 15 oz, or what is observed about  $\bar{Y}$  supports the suspicion that the average weight of a can of peaches is below 15 oz.

### Ingredients of a significance test

1. Aspect of a population that is of interest.
2. Null hypothesis  $H_0$ .
3. Alternative hypothesis  $H_A$ .
4. Statistic and null distribution.
5. Measure evidence against  $H_0$ .
6. Make a decision and interpret in the context of the problem.

## Significance Tests with Known $\sigma$

### Ingredients of a significance test

1. Of interest is the population mean  $\mu$ .
2. The null hypothesis is

$$H_0 : \mu = 15.$$

Remark:  $H_0$  is the claim initially favored or believed to be true. Thus the results of the sample reflect random variation.

3. The alternative hypothesis is

$$H_A : \mu < 15.$$

Remark:  $H_A$  is the departure from  $H_0$  that one wishes to be able to detect. Thus the results of the sample reflect some factor other than random variation.

4. The statistic is the sample mean  $\bar{Y}$  based on a random sample of size  $n = 60$  ( $Y_1, Y_2, \dots, Y_n$ ). Under  $H_0$ ,

$$\bar{Y} \sim N\left(15, \frac{(0.3)^2}{60}\right).$$

That is,

$$\bar{Y} \sim N(15, (0.0387)^2).$$

Remark: The null distribution is based on FACT 1.

## Significance Tests with Known $\sigma$

### Ingredients of a significance test

5. Suppose the observed sample mean is

$$\bar{y} = 14.9$$

based on an observed sample of size  $n = 60$  ( $y_1, y_2, \dots, y_n$ ). Is this consistent with the claim or the suspicion? We measure the evidence against  $H_0$  by computing a *p-value*, which is the probability of having data as extreme or more extreme than the observed data, if  $H_0$  were true. For the given sample,

$$\begin{aligned} \text{p-value} &= P(\bar{Y} \leq 14.9 | H_0) \\ &= P\left(\frac{\bar{Y} - 15}{0.3/\sqrt{60}} \leq \frac{14.9 - 15}{0.3/\sqrt{60}}\right) \\ &= P(Z \leq -2.58) \\ &= 0.0049 \end{aligned}$$

6. Either  $H_0$  is true but we have observed an event with probability 0.0049, which is very rare, or  $H_0$  is false.

## Significance Tests with Known $\sigma$

### Remarks

- Alternatively, use:

$$Z = \frac{\bar{Y} - 15}{0.0387}.$$

Under  $H_0$ ,  $\bar{Y} \sim N(15, (0.0387)^2)$  and thus  $Z \sim N(0, 1)$ .

- The observed z-score is

$$z = \frac{14.9 - 15}{0.0387} = -2.58.$$

- The p-value is

$$P(Z \leq -2.58) = 0.0049.$$

- The statistic  $Z = \frac{\bar{Y}-15}{0.0387}$  here is a *test statistic*, which refers to a quantity that one can compute from the sample and can be directly interpreted using a statistical table.

## Significance Tests with Known $\sigma$

### Remarks

- Often the p-value is compared against  $\alpha = 0.05$ . If p-value  $< 0.05$ , then we say

“Reject  $H_0$  at the 5% level”

or

“The results are significant at the 5% level”

- The 5% level is known as the  $\alpha$ -level. We can have different  $\alpha$ -levels. Some common  $\alpha$ -levels are 1%, 5%, and 10%.
- The p-value can be interpreted as evidence against  $H_0$ . The smaller the p-value, the greater the evidence. Roughly, we can interpret a p-value
  - larger than 0.10 as no evidence against  $H_0$ ;
  - between 0.05 and 0.10 as weak evidence against  $H_0$ ;
  - between 0.01 and 0.05 as moderate evidence against  $H_0$ ;
  - between 0.001 and 0.01 as strong evidence against  $H_0$ ;
  - smaller than 0.001 as very strong evidence against  $H_0$ .
- $H_0$  is either true or not true. Thus p-value is not  $P(H_0 \text{ is true})$ .

## Significance Tests with Known $\sigma$

### Remarks

- In the fruit can example, an observed sample mean  $\bar{y}$  close to 15 oz relative to the standard error would result in:
  - a large p-value;
  - do not reject  $H_0$  or accept  $H_0$ ;
  - finding  $H_0$  plausible;
  - the data supporting the claim.

On the other hand, an observed sample mean  $\bar{y}$  far below 15 oz relative to the standard error would result in:

- a small p-value;
- reject  $H_0$  or accept  $H_A$ ;
- finding weak/moderate/strong/very strong evidence against  $H_0$ ;
- the data not supporting the claim.

## Significance Tests with Known $\sigma$

### Remarks

- In general, what we wish to establish is set up as the  $H_A$ . For example:

$$H_0 : \mu = 15, \quad H_A : \mu < 15$$

If  $H_0$  is rejected, then the suspicion is plausible. Another example:

$$H_0 : \mu = 14.9, \quad H_A : \mu > 14.9$$

If  $H_0$  is rejected, then the claim is plausible.

- Significance testing also needs to follow rules to be valid mathematically. Here we focus on

$$H_0 : \mu = \mu_0$$

for a given  $\mu_0$  and three types of  $H_A$

$$H_A : \mu > \mu_0, \quad H_A : \mu < \mu_0, \quad H_A : \mu \neq \mu_0$$

## Significance Tests with Known $\sigma$

### Tomato weight example

A given variety of tomatoes have a mean weight of 30 gm and weights are approximately normal with  $\sigma^2 = 5^2$ . A new fertilizer is introduced and its use may affect the mean weight of tomatoes (but not the variance). Sixteen (16) tomatoes which have been grown with the new fertilizer are weighed and a sample mean  $\bar{Y}$  is calculated. We wish to determine whether there is evidence of an effect of the new fertilizer on the mean weight of the tomatoes.

## Significance Tests with Known $\sigma$

### Tomato weight example

1. Of interest is  $\mu =$  population mean weight of this given variety of tomatoes.
2.  $H_0 : \mu = 30$  (i.e., there is no fertilize effect).
3.  $H_A : \mu \neq 30$  (i.e., there is a fertilize effect).
4. Under  $H_0$ ,

$$\bar{Y} \sim N\left(30, \frac{5^2}{16}\right)$$

(i.e.,  $\mu_{\bar{Y}} = 30, \sigma_{\bar{Y}} = 1.25$ ). Or the test statistic is

$$Z = \frac{\bar{Y} - 30}{1.25} \sim N(0, 1).$$

5. Suppose the observed sample mean is  $\bar{y} = 32.5$ . Then

$$z = \frac{32.5 - 30}{1.25} = 2$$

and the p-value is

$$\begin{aligned} \text{p-value} &= P(Z \leq -2) + P(Z \geq 2) = 2 \times P(Z \geq 2) \\ &= 2 \times 0.0228 = 0.0456. \end{aligned}$$

6. Reject  $H_0$  at the 5% level. There is weak/moderate evidence against  $H_0$ .

## Significance Tests with Known $\sigma$

### Remarks

- Here  $H_A$  is a two-sided alternative.
- The observed  $\bar{y}$  is 2 standard deviation above the population mean under  $H_0$  (i.e.  $z = 2$ ). If the observed  $\bar{y}$  were 2 standard deviation below the population mean (i.e.  $z = -2$ ), we would be equally surprised.
- That is, if the observed  $\bar{y}$  is 32.5 or larger provides evidence against  $H_0$ , then so does an observed  $\bar{y}$  of 27.5 or smaller.
- Decide on a one-sided versus a two-sided  $H_A$  before looking at the data.
- In general, use a two-sided  $H_A$  unless there are strong reasons to use a one-sided  $H_A$ .
- In practice, the population variance  $\sigma^2$  is not known. Thus the term

$$Z = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}}$$

cannot be calculated. We estimate  $\sigma^2$  by the sample variance  $S^2$ .

## Significance Tests with Unknown $\sigma$

### Main idea

- Consider a random sample  $Y_1, Y_2, \dots, Y_n$  of size  $n$  from normal distribution  $N(\mu, \sigma^2)$  with unknown  $\mu$  and unknown  $\sigma^2$ .
- Suppose the significance test of interest is:

$$H_0 : \mu = \mu_0$$

for a given  $\mu_0$ .

- The following is the test statistic

$$T = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}}$$

- A very useful fact is that, under  $H_0$ ,

$$T \sim T_{n-1}$$

where  $T_{n-1}$  is a *T distribution* with  $n - 1$  degrees of freedom (d.f.).

## Significance Tests with Unknown $\sigma$

### T distribution

- Like a standard normal distribution  $Z$ , a T-distribution is also defined by a bell-shaped distribution curve and is symmetric about 0.
- But a T-distribution has heavier tails than the  $Z$  distribution curve.
- The shape of the distribution curve depends on the sample size  $n$ . For larger  $n$ , the distribution curve has thinner tails, less spread, and more  $Z$ -like.
- In fact,  $T_\infty = Z$ .
- The d.f. are the same as the d.f. of the sample variance  $S^2$ .

## Significance Tests with Unknown $\sigma$

### Table C

- Table C on page 410 of the bluebook gives

$$P(T \geq a)$$

for various d.f. (row index). The selected  $a$  values are inside of Table C and the corresponding probabilities are on the top of Table C (column index).

	df	...	0.10	0.05	0.025	...	
	1						
	2						
	...						
	9	...	1.833		...		
	...						

- For example,

means  $P(T_9 \geq 1.833) = 0.05$ .

- Table C gives both upper tail probabilities such as

$$P(T_9 \geq 1.833) = 0.05$$

and quantiles such as the 0.95<sup>th</sup> quantile  $q_{0.95} = 1.833$

$$P(T_9 \geq q_{0.95}) = 0.05.$$

## Significance Tests with Unknown $\sigma$

### Wheat yield example

Six (6) 1-acre plots are sown with a new variety of wheat. The yields are (in cut/acre):

25, 21, 24, 20, 26, 22

and are from  $N(\mu, \sigma^2)$ . Is there evidence that the population mean yield for this variety of wheat differs from 20 cut/acre?

## Significance Tests with Unknown $\sigma$

### Wheat yield example

- Let  $\mu$  = population mean yield of this variety of wheat.
- $H_0 : \mu = 20$  versus  $H_A : \mu \neq 20$  (i.e.  $\mu_0 = 20$ ).
- Under  $H_0$ , the test statistic is

$$T = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}} \sim T_{n-1}$$

- Here  $n = 6, \bar{y} = 23, s = 2.37$ .
- The observed test statistic is

$$t = \frac{\bar{y} - \mu_0}{s/\sqrt{n}} = \frac{23 - 20}{2.37/\sqrt{6}} = 3.10.$$

on d.f. =  $n - 1 = 5$ .

- Thus the p-value is

$$\text{p-value} = 2 \times P(T_5 \geq 3.10).$$

Since  $0.01 < P(T_5 \geq 3.10) < 0.025$  from Table C, the p-value is between 0.02 and 0.05.

- Reject  $H_0$  at the 5% level, but not at the 1% level. There is moderate evidence against  $H_0$ .

## Significance Tests with Unknown $\sigma$

### Remarks

- An observed sample mean  $\bar{y}$  close to 20 relative to the standard error would result in:
  - a small  $|t|$  and a large p-value;
  - do not reject  $H_0$  or accept  $H_0$ ;
  - finding  $H_0$  plausible;
  - the data supporting  $H_0$ .

On the other hand, an observed sample mean  $\bar{y}$  far from 20 relative to the standard error would result in:

- a large  $t$  and a small p-value;
  - reject  $H_0$  or accept  $H_A$ ;
  - finding weak/moderate/strong/very strong evidence against  $H_0$ ;
  - the data not supporting  $H_0$ .
- This is known as a t-test.

## Significance Tests with Unknown $\sigma$

### Milk yield example

A farmer wishes to sell a herd in Wisconsin and claims that the mean milk yield of cows in this population is at least 40 lbs/day. But there is suspicion about this claim. Consider milk yield in lb/day for fourteen (14) cows from this herd on a given day. Assume that data form a random sample of size 14 from normal distribution  $N(\mu, \sigma^2)$ .

## Significance Tests with Unknown $\sigma$

### Milk yield example

- Let  $\mu$  = population mean milk yield of this herd.
- $H_0 : \mu \geq 40$  versus  $H_A : \mu < 40$  (i.e.  $\mu_0 = 40$ ).
- Under  $H_0$ , the test statistic is

$$T = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}} \sim T_{n-1}$$

- Suppose the observed  $\bar{y} = 36.2$ ,  $s^2 = 95.26$ . Then the observed test statistic is:

$$t = \frac{\bar{y} - \mu_0}{s/\sqrt{n}} = \frac{36.2 - 40}{\sqrt{95.26/14}} = -1.46$$

on d.f. =  $n - 1 = 13$ .

- Thus the p-value is

$$p - value = P(T_{13} \leq -1.46),$$

which is between 0.05 and 0.10.

- Reject  $H_0$  at the 10% level but do not reject  $H_0$  at the 5% level. There is weak evidence against  $H_0$ .

## Significance Tests with Unknown $\sigma$

### Remarks

- If  $H_0 : \mu = 40$  is rejected, then any  $\mu_0$  above 40 will also be rejected.
- Again the issue of one-sided alternative versus two-sided alternative as for significance tests with known  $\sigma$ .
- Assumptions on a t-test:
  1.  $Y_1, \dots, Y_n$  form a random sample of size  $n$  from  $N(\mu, \sigma^2)$ .
  2.  $\sigma^2$  is unknown (i.e., no assumption).
  3. any sample size  $n$  (i.e., no assumption).
- T-test is robust for non-normality, but is very sensitive to dependence.

## Significance Tests with Binomial Data

### Coin toss example

Toss a coin 100 times independently. Let  $Y = \#$  of heads and  $Y \sim B(100, p)$ . Perform a significance test on whether  $p$  is 0.65 or not.

- The parameter of interest is the probability of heads  $p$ .
- $H_0 : p = 0.65$  versus  $H_A : p \neq 0.65$ .
- Under  $H_0$ , we have  $Y \sim B(100, 0.65)$  with  $\mu_Y = 65$ ,  $\sigma_Y^2 = 22.75 = (4.77)^2$ .
- Suppose the observed  $\#$  of heads is  $y = 74$ . Then the p-value is  $P(Y \geq 74) + P(Y \leq 56)$
- To compute the p-value, we use normal approximation with  $Y_{\text{NA}} \sim N(65, (4.77)^2)$ . Thus  $Z = \frac{Y_{\text{NA}} - 65}{4.77} \sim N(0, 1)$  and the approximate p-value is

$$\begin{aligned} P(Y \geq 74) + P(Y \leq 56) &= P(Y_{\text{NA}} \geq 74) + P(Y_{\text{NA}} \leq 56) \\ &= P(Z \geq 1.89) + P(Z \leq -1.89) \\ &= 2 \times 0.0294 = 0.0588. \end{aligned}$$

- Reject  $H_0$  at the 5% level, but not at the 10% level. There is weak evidence against  $H_0$ .

## Significance Tests with Binomial Data

### Remarks

- We could have used binomial distribution directly to compute an exact p-value, but hand calculation is tedious.

$$P(Y \geq 74) + P(Y \leq 56) = p(0) + \cdots + p(56) + p(74) + \cdots + p(100) = 0.0741.$$

- We could have used the proportion of heads as the test statistic

$$\hat{p} = \frac{Y}{n}$$

with an approximate normal distribution

$$\hat{p}_{\text{NA}} \sim N\left(p, \frac{pq}{n}\right).$$

- In the coin toss example,  $\hat{p}_{\text{NA}} \sim N(0.65, (0.0476)^2)$  under  $H_0$ . Thus

$$Z = \frac{\hat{p}_{\text{NA}} - 0.65}{0.0476} \sim N(0, 1).$$

Since the observed  $\hat{p} = 74/100$ , the p-value is

$$\begin{aligned} 2 \times P(\hat{p} \geq 0.74) &\approx 2 \times P\left(Z \geq \frac{0.74 - 0.65}{0.0476}\right) \\ &= 2 \times P(Z \geq 1.89) = 0.0588. \end{aligned}$$

Same p-value as before!

## Significance Tests with Binomial Data

### Remarks

- We can use the normal approximation, because  $np = 65$ ,  $nq = 35$  are both larger than 5.
- If  $np < 5$  or  $nq < 5$ , then we cannot use normal approximation.
- For example, for  $Y \sim B(10, p)$ , we wish to test:

$$H_0 : p = 0.2 \text{ versus } H_A : p < 0.2.$$

Under  $H_0$ ,  $Y \sim B(10, 0.2)$ . Since  $np = 2 < 5$ , we cannot use normal approximation. Instead use binomial distribution directly.

- Suppose the observed  $y = 0$ . Then the p-value is

$$P(Y \leq 0) = (0.8)^{10} = 0.107.$$

Thus there is no/weak evidence against  $H_0$ .

- If  $H_A : p \neq 0.2$ , then we define p-value as

$$2 \times P(Y \leq 0) = 2 \times 0.107 = 0.214.$$

- Note that there is no  $\bar{Y}$  involved. The r.v. of interest is  $Y \sim B(n, p)$ .

# Significance Tests Using CLT

## Main idea

- Consider a random sample  $Y_1, Y_2, \dots, Y_n$  of size  $n$  from an arbitrary distribution  $D(\mu, \sigma^2)$  with unknown  $\mu$  and unknown  $\sigma^2$ .

- Recall the CLT,

$$\bar{Y} \approx N\left(\mu, \frac{\sigma^2}{n}\right).$$

- Since for large  $n$ ,  $S^2 \approx \sigma^2$ , a useful fact is that

$$\frac{\bar{Y} - \mu}{S/\sqrt{n}} \approx N(0, 1).$$

- Suppose the significance test of interest is:

$$H_0 : \mu = \mu_0$$

for a given  $\mu_0$ .

- The following test statistic

$$\frac{\bar{Y} - \mu_0}{S/\sqrt{n}}$$

is approximately  $N(0, 1)$ .

## Significance Tests Using CLT

### Remarks

- If the distribution is normal, then regardless of the sample size  $n$ ,

$$\frac{\bar{Y} - \mu_0}{S/\sqrt{n}} \sim T_{n-1}$$

and the distribution is exact, not approximate.

- If  $n$  is large, then  $T_{n-1}$  is more and more like  $N(0, 1)$ . Thus the new fact here is consistent with the previous results.

## Significance Tests Using CLT

### Check weight example

A scale is used to weigh a check weight. Over the years, the weight is 40.6 gm. After a lab fire, the scale is re-tested to see if it is affected by the fire. One hundred (100) measures are taken and the observed  $\bar{y} = 41.6$ ,  $s = 5$ .

- Let  $\mu =$  mean weight of the check weight after the fire.
- $H_0 : \mu = 40.6$  versus  $H_A : \mu \neq 40.6$  (i.e.  $\mu_0 = 40.6$ ).
- By the CLT, under  $H_0$ ,

$$Z = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}} \approx N(0, 1).$$

- Thus the observed test statistic is

$$z = \frac{41.6 - 40.6}{5/\sqrt{100}} = 2.0.$$

and the p-value is

$$2 \times P(Z \geq 2.0) = 2 \times 0.0228 = 0.0456.$$

- Reject  $H_0$  at the 5% level, but not the 1% level. There is weak/moderate evidence against  $H_0$ .

# Significance Tests

## A quick summary

$Y_1, \dots, Y_n$	$\sigma^2$	test on	test stats	test is
$N(\mu, \sigma^2)$	known	$\mu$	$Z = \frac{Y - \mu}{\sigma/\sqrt{n}}$ $\sim N(0, 1)$	exact
$D(\mu, \sigma^2)$	known $n$ is large	$\mu$	$Z = \frac{Y - \mu}{\sigma/\sqrt{n}}$ $\sim N(0, 1)$	approximate
$N(\mu, \sigma^2)$	unknown	$\mu$	$Z = \frac{Y - \mu}{S/\sqrt{n}}$ $\sim T_{n-1}$	exact
$D(\mu, \sigma^2)$	unknown $n$ is large	$\mu$	$Z = \frac{Y - \mu}{S/\sqrt{n}}$ $\sim N(0, 1)$	approximate
$B(n, p)$	$np \geq 5$ $nq \geq 5$	$p$	$Z = \frac{Y_{NA} - \mu_Y}{\sigma_Y}$ $\sim N(0, 1)$ $\mu_Y = np$ $\sigma_Y^2 = npq$	approximate

## Significance Tests

Another quick summary

$Y_1, \dots, Y_n$	$n$	$\sigma$ known	$\sigma$ unknown
$N(\mu, \sigma^2)$	small	$Z$	$T$
$N(\mu, \sigma^2)$	large	$Z$	$T$
$D(\mu, \sigma^2)$	small	no general result	no general result
$D(\mu, \sigma^2)$	large	$Z$ by CLT	$Z$ by CLT

## Significance Tests

### Another look at the $\alpha$ -level

- Recall that  $\alpha$  is the level of significance testing. For example,  $\alpha = 0.01, 0.05, 0.10$ .
- In significance testing, we can make two types of mistakes:
  - Type I error: reject  $H_0$  when  $H_0$  is true.
  - Type II error: accept  $H_0$  when  $H_A$  is true.
- Then the probability of type I error is  $\alpha$ .
- In the milk yield example, under  $H_0 : \mu = 40$ , from Table C,
$$\alpha = 0.05 = P(T_{13} \leq -1.771 | H_0)$$
- The observed  $t = -1.46$  and the p-value is between 0.05 and 0.10.
- Do not reject  $H_0$  at the 5% level, because the observed  $t$  is above -1.771.
- This is another way of thinking about significance testing.
- We will return to this later on.

# Significance Tests

## Key R commands

```
> #fruit can example p-value
> pnorm(14.9, mean=15, sd=0.3/sqrt(60), lower.tail=T)
[1] 0.004911637
```

```
> #tomato weight example p-value
> 2*pnorm(32.5, mean=30, sd=5/sqrt(16), lower.tail=F)
[1] 0.04550026
```

```
> #wheat yield example
> wheat = c(25, 21, 24, 20, 26, 22)
> t.test(wheat, mu=20, alternative="two.sided")
```

One Sample t-test

```
data: wheat
t = 3.1053, df = 5, p-value = 0.02669
alternative hypothesis: true mean is not equal to 20
95 percent confidence interval:
 20.51658 25.48342
sample estimates:
mean of x
      23
```

```
> #alternatively
> wheat.ybar = mean(wheat)
> wheat.s = sd(wheat)
> wheat.n = length(wheat)
> wheat.mu = 20
> wheat.t = (wheat.ybar-wheat.mu)/(wheat.s/sqrt(wheat.n))
> wheat.t
[1] 3.105295
> 2*pt(wheat.t, wheat.n-1, lower.tail=F)
[1] 0.02669259
```

# Significance Tests

## Key R commands

```
> #milk yield example
> milk.ybar = 36.2
> milk.s = sqrt(95.26)
> milk.n = 14
> milk.mu = 40
> milk.t = (milk.ybar-milk.mu)/(milk.s/sqrt(milk.n))
> milk.t
[1] -1.456774
> pt(milk.t, milk.n-1, lower.tail=T)
[1] 0.08445275

> #coin toss example
> #exact p-value
> pbinom(56, size=100, prob=0.65, lower.tail=T)+pbinom(73, size=100, prob=0.65, lower.tail=F)
[1] 0.07405748
> #normal approximation
> 2*pnorm(74, mean=100*0.65, sd=sqrt(100*0.65*0.35),lower.tail=F)
[1] 0.05917207
> #success proportion
> 2*pnorm(74/100, mean=0.65, sd=sqrt(0.65*0.35/100),lower.tail=F)
[1] 0.05917207

> #check weight example
> 2*pnorm(41.6, mean=40.6, sd=5/sqrt(100), lower.tail=F)
[1] 0.04550026
```