Ch 8: Inference for two samples

Contents

1 Preliminaries **2**
1.1 Prelim: Two Normals . . . 3
1.2 Two sample z-test . . . 15
1.3 Pooled T-test . . . 39
1.4 Paired T-test . . . 43
1.5 Two Sample Inference on Proportion . . . 48
1.6 Two Sample Inference on Variance . . . 55

July 10, 2017

1

Preliminaries

[ToC]

1.1 Prelim: Two Normals

[ToC]

Suppose

$$X \sim N(\mu_1, \sigma_1) \qquad \text{and} \qquad Y \sim N(\mu_2, \sigma_2),$$

X and Y are independent. How do you calculate

$$P(X < Y) =?$$

Answer

Since any linear combination of two normals are also normal, we have

$$X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

Then we can calculate the probability as

$$P(X < Y) = P(X - Y < 0).$$

Example:

Suppose X is a diameter of a shaft, which is $N(13, .4)$. Y is a diameter of a shaft cover, which is $N(13.2, .3)$. For a proper fit, a shaft is supposed to fit into a shaft cover with a clearance between .1 and .3. What is the probability of proper fit?

Example:

Suppose X is a diameter of a shaft, which is $N(13, .4)$. Y is a diameter of a shaft cover, which is $N(13.2, .3)$. For a proper fit, a shaft is supposed to fit into a shaft cover with a clearance between .1 and .3. What is the probability of proper fit?

$$P(\text{ proper fit}) = P(.1 \leq Y - X \leq .3)$$

Example:

Suppose X is a diameter of a shaft, which is $N(13, .4)$. Y is a diameter of a shaft cover, which is $N(13.2, .3)$. For a proper fit, a shaft is supposed to fit into a shaft cover with a clearance between .1 and .3. What is the probability of proper fit?

$$P(\text{ proper fit}) = P(.1 \leq Y - X \leq .3)$$

I know that

$$Y - X \sim N(\mu_*, \sigma_*^2).$$

Need to calculate μ_* and σ_*.

Example Cont'd

Mean of $Y - X$ is

$$\mu_* = E(Y - X) = E(Y) - E(X) = 13.2 - 13 = .2$$

Example Cont'd

Mean of $Y - X$ is

$$\mu_* = E(Y - X) = E(Y) - E(X) = 13.2 - 13 = .2$$

Variance of $Y - X$ is

$$\sigma_*^2 = V(Y - X) = V(Y) + V(X) = .4 + .3 = .7$$

Example Cont'd

Mean of $Y - X$ is

$$\mu_* = E(Y - X) = E(Y) - E(X) = 13.2 - 13 = .2$$

Variance of $Y - X$ is

$$\sigma_*^2 = V(Y - X) = V(Y) + V(X) = .4 + .3 = .7$$

$$Y - X \sim N(.2, .7).$$

Example Cont'd

$$P(\text{ proper fit}) = P(.1 \leq Y - X \leq .3)$$

Example Cont'd

$$
\begin{aligned}
P(\text{ proper fit}) &= P(.1 \leq Y - X \leq .3) \\
&= P(Y - X \leq .3) - P(Y - X \leq .1)
\end{aligned}
$$

Example Cont'd

$$
\begin{aligned}
P(\text{ proper fit}) &= P(.1 \leq Y - X \leq .3) \\
&= P(Y - X \leq .3) - P(Y - X \leq .1)
\end{aligned}
$$

Since $Y - X \sim N(.2, .7)$,

$$
= \Phi\Big(\frac{.3 - .2}{\sqrt{.7}}\Big) - \Phi\Big(\frac{.1 - .2}{\sqrt{.7}}\Big)
$$

Example Cont'd

$$
\begin{aligned}
P(\text{ proper fit}) &= P(.1 \leq Y - X \leq .3) \\
&= P(Y - X \leq .3) - P(Y - X \leq .1)
\end{aligned}
$$

Since $Y - X \sim N(.2, .7)$,

$$
\begin{aligned}
&= \Phi\left(\frac{.3 - .2}{\sqrt{.7}}\right) - \Phi\left(\frac{.1 - .2}{\sqrt{.7}}\right) \\
&= \Phi(.12) - \Phi(-.12) \\
&= 0.096
\end{aligned}
$$

1.2 Two sample z-test

Let's use the same principle, and get inference on two-sample problem.

One-sample inference: There's population with true mean μ. Want to know μ.

- Estimate μ $\Rightarrow$ CI
- Is $\mu > \mu_0$? $\Rightarrow$ Test of hypothesis (z-test, t-test)

Two-sample inference: There's population 1 with true mean μ_1.
There's population 2 with true mean μ_2.

- Can we estimate the difference $\mu_1 - \mu_2$? $\Rightarrow$ CI
- Is $\mu_1 = \mu_2$? Is $\mu_1 > \mu_2 + 10$? μ $\Rightarrow$ Test of hypothesis

Two-sample

- Assume population 1 has $N(\mu_1, \sigma_1^2)$ distribution. Then draw sample $X_1, \ldots, X_{n_1}$ from population 1.

$$\text{Calculate} \quad \bar{X}, \quad S_1. \quad \text{Sample size} = n_1.$$

- Assume population 2 has $N(\mu_2, \sigma_2^2)$ distribution. Then draw sample $Y_1, \ldots, Y_{n_2}$ from population 2.

$$\text{Calculate} \quad \bar{Y}, \quad S_2. \quad \text{Sample size} = n_2.$$

Example:

- aaa
- aaa

Two sample test for difference in mean

How can we use $\bar{X}$, and $\bar{Y}$ and test hypothesis

$$H_0 : \mu_1 = \mu_2$$

$$\begin{aligned} H_A : \mu_1 &> \mu_2 \text{ (upper-tailed)} \\ \mu_1 &< \mu_2 \text{ (lower-tailed)} \\ \mu_1 &\neq \mu_2 \text{ (two-tailed)} \end{aligned}$$

Two Normals

We know that

$$\bar{X} \sim N\Big(\mu_1, \frac{\sigma_1^2}{n_1}\Big) \qquad \text{and} \qquad \bar{Y} \sim N\Big(\mu_2, \frac{\sigma_2^2}{n_2}\Big).$$

Then, since they are independent,

$$\bar{X} - \bar{Y} \sim N\Big(\mu_1 - \mu_2, \quad \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\Big)$$

This looks kinda like in one-sample case,

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

Analogy

one-sample case	two-sample case
$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$	$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \quad \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}\right)$
$H_0 : \mu = \mu_0$	$H_0 : \mu_1 = \mu_2$
$H_A : \mu > \mu_0$	$H_A : \mu_1 > \mu_2$
$\mu < \mu_0$	$\mu_1 < \mu_2$
$\mu \neq \mu_0$	$\mu_1 \neq \mu_2$

We need to rewrite the hypothesis

Analogy

one-sample case	two-sample case	if $\Delta_0 = 0$
$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$	$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \quad \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}\right)$	
$H_0 : \mu = \mu_0$	$H_0 : \mu_1 - \mu_2 = \Delta_0$	$(\mu_1 = \mu_2)$
$H_A : \mu > \mu_0$	$H_A : \mu_1 - \mu_2 = \Delta_0$	$(\mu_1 > \mu_2)$
$\mu < \mu_0$	$\mu_1 - \mu_2 < \Delta_0$	$(\mu_1 < \mu_2)$
$\mu \neq \mu_0$	$\mu_1 - \mu_2 \neq \Delta_0$	$(\mu_1 \neq \mu_2)$
$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$z = \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	

Two sample z-test To test the null hypothesis of

$$H_0 : \mu_1 - \mu_2 = \Delta_0$$

against alternatives

$$\begin{aligned} H_A : \mu_1 - \mu_2 &> \Delta_0 \quad \text{(Upper-tailed alternative)} \\ \mu_1 - \mu_2 &< \Delta_0 \quad \text{(Lower-tailed alternative)} \\ \mu_1 - \mu_2 &\neq \Delta_0 \quad \text{(Two-tailed alternative)}\ , \end{aligned}$$

we use the test statistic

$$z = \frac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

and rest is the same old z-test:

H_A	rejection region	p-value	power
upper-tailed	$z > z_\alpha$	$1 - \Phi(z)$	$1 - \Phi(z_\alpha - \mu_A^*)$
lower-tailed	$z < -z_\alpha$	$\Phi(z)$	$\Phi(-z_\alpha - \mu_A^*)$
Two-tailed	$z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$	$2\big(1 - \Phi(\lvert z\rvert)\big)$	$1 - \Phi(z_{\frac{\alpha}{2}} - \mu_A^*) + \Phi(-z_{\frac{\alpha}{2}} - \mu_A^*)$

$$\mu_A^* = \frac{\mu_1 - \mu_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Two sample CI for $\mu_1 - \mu_2$

one-sample case	two-sample case
$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$	$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \quad \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}\right)$
CI for μ	CI for $\mu_1 - \mu_2$
$\bar{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$	$(\overline{X} - \overline{Y}) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Example: Tire and Fuel

(Walpole) p359 A taxi company manager is trying to decide whether the use of radial tires instead of regular belted tires improves fuel economy. Twelve cars were equipped with radial tires and driven over a prescribed test course. Without changing drivers, the same cars were then equipped with regular belted tires and driven once again over the test course. The gasoline consumption, in kilometers per liter, was recorded as follows:

Can we conclude that cars equipped with radial tires give better fuel economy than those equipped with belted tires? Assume the populations to be normally distributed.

Assume that the observations are random sample from the Normal distributions, and $\sigma_1 = \sigma_2 = 1$

Data: (tire and fuel)

```
    Car Radial Belted  (Km/l)
D=c(1 ,   4.2,    4.1,
    2 ,   4.7,    4.4,
    3 ,   6.6,    6.4,
    4 ,   7.0,    6.7,
    5 ,   6.7,    6.4,
    6 ,   4.5,    4.4,
    7 ,   5.7,    5.7,
    8 ,   6.0,    5.8,
    9 ,   7.4,    6.5,
    10,   4.9,    4.7,
    11,   6.1,    6.0,
    12,   5.2,    4.9)

dim(D) <- c(3,12)
D <- t(D)

t.test(D[,2], D[,3])
```

$\bar{X} = 5.75 \quad \bar{Y} = 5.50$ Is this an evidence that $\mu_1 > \mu_2$?

$S_1^2 = 1.053 \quad S_2^2 = 0.945$

Analysis

Assuming $\sigma_1^2 = \sigma_2^2 = 1$, we want to test

$$\begin{aligned} H_0: &\quad \mu_1 - \mu_2 = 0 \\ H_A: &\quad \mu_1 - \mu_2 > 0 \quad (upper - tailed) \end{aligned}$$

Test Statistic

$$z = \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Analysis

Assuming $\sigma_1^2 = \sigma_2^2 = 1$, we want to test

$$\begin{aligned} H_0: &\quad \mu_1 - \mu_2 = 0 \\ H_A: &\quad \mu_1 - \mu_2 > 0 \quad (upper - tailed) \end{aligned}$$

Test Statistic

$$z = \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{5.75 - 5.50 - 0}{\sqrt{\frac{1}{12} + \frac{1}{12}}} = .6124$$

P-value $= 1 - \Phi(.6124) = .27$.
Can't reject H_0 for $\alpha = .05$. $\Rightarrow$ Not enough evidence to claim Radial tire has higher milage on average.

(Difference is not significant)

When S_1, S_2 must be used instead of σ_1, σ_2.

One-sample case,

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \qquad \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$$

That's why z-test became t-test.

Two-sample case,

$$\frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1) \qquad \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t(\nu)$$

Only difference of the degrees of freedom ν.

Degrees of freedom

The degrees of freedom ν can be calculated by the formula

$$\nu = \frac{(a+b)^2}{\frac{a^2}{n_1-1} + \frac{b^2}{n_2-1}} \qquad \text{where} \qquad a = \frac{S_1^2}{n_1}, \qquad b = \frac{S_2^2}{n_2}.$$

If you are in hurry, you can use $\min(n_1, n_2) - 1$.

Example:

Two new methods for producing a tire have been proposed. To ascertain which is superior, a tire manufacturer produces a sample of 10 tires using the first method and a sample of 8 using the second. The first set is to be road tested at location A and the second at location B.

It is known that the lifetimes of tires tested at location A or B are normally distributed.

Manufacturer is interested in testing the hypothesis that there is no appreciable difference in the mean life of tires produced by either method.

Data:

```
Tire Lives in Units of 100 Kilometers

Tires_at_A      Tires_at_B
61.1            62.2
58.2            56.6
62.3            66.4
64              56.2
59.7            57.4
66.2            58.4
57.8            57.6
61.4            65.4
62.2
63.6

X <- c(61.1, 58.2, 62.3, 64, 59.7, 66.2, 57.8, 61.4, 62.2, 63.6);
Y <- c(62.2, 56.6, 66.4, 56.2, 57.4, 58.4, 57.6, 65.4);

t.test(X,Y)
```

$\bar{X} = 61.65 \quad \bar{Y} = 60.03$ $\leftarrow$ Is this an meaningful difference?

$S_1^2 = 2.62 \quad S_2^2 = 4.07.$

Analysis

We want to test

$$\begin{aligned} H_0: &\quad \mu_1 - \mu_2 = 0 \\ H_A: &\quad \mu_1 - \mu_2 > 0 \quad (upper - tailed) \end{aligned}$$

Test Statistic

$$z = \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{61.65 - 60.03 - 0}{\sqrt{\frac{1}{10} + \frac{1}{8}}} = .979$$

P-value $= 1 - \Phi(.6124) = .27$.
Can't reject H_0 for $\alpha = .05$. $\Rightarrow$ Not enough evidence to claim Radial tire has higher milage on average.

(Difference is not significant)

Example: Arizona Water

(Montogomery p.342) Arsenic concentration in public drinking water supplies is a potential health risk. An article in the Arizona Republic (Sunday, May 27, 2001) reported drinking water arsenic concentrations in parts per billion (ppb) for 10 methropolitan Phoenix communities and 10 communities in rural Arizona.

```
Metro  Phoenix          Rural  Arizona
Phoenix        3        Rimrock        48
Chandler       7        Goodyear       44
Gilbert       25        New_River      40
Glendale      10        Apachie_Jtn    38
Mesa          15        Buckeye        33
Paradise_Vly   6        Nogales        21
Peoria        12        Black_Canyon   20
Scottsdale    25        Sedona         12
Tempe         15        Payson          1
Sun_City       7        Casa_Grande    18
```

```
A=c(3,48,  7,   44, 25,   40,10,   38,15,   33,6,   21,12,   20,25,   12,15,    1,7,   18)
D=matrix(A, 10,2,byrow=T)

t.test(D[,1], D[,2])
```

Example: Battery Life

- Duracell Alkaline AA batteries vs Eveready Energizer Alkaline AA batteries. 4.5 hours and 4.2 hours, respectively.
- both sample size are 150
- The population standard deviations of lifetime are 1.8 hours for Duracell and 2.0 hours for Eveready batteries.

Test, with significance level of .05, the hypothesis of true mean lifetime of Duracell batteries is longer than that of Eveready brand.

```
x <- 4.5
y <- 4.2
m <- 150
n <- 150
s1 <- 1.8
s2 <- 2
(x-y - 0) / sqrt( s1^2/m + s2^2/n )

1 - pnorm( 1.96 - (.5 - 0) / sqrt( s1^2/m + s2^2/n ))
```

Example:

- Stopping distances from 50 mph for two different types of braking systems.
- System A: $n_1 = 6$, $\overline{X} = 76$, $S_1 = 5$,
- System B: $n_2 = 6$, $\overline{Y} = 88$, $S_2 = 5.5$.

Use the two-sample $t-$ test at significance level .01 to see if the data showes evidence that the population mean stopping distance of the breaking system A is more than 10 meter shorter than that of system B.

```
x <- 76
y <- 95
m <- 16
n <- 20
s1 <- 12.3
s2 <- 11.5

(x-y - (-10)) / sqrt( s1^2/m + s2^2/n )

pt( -2.01 - (-20 - (-10) ) / sqrt( s1^2/m + s2^2/n ), 15)
```

1.3 Pooled T-test

Test Statistic for t-test

$$z = \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

S_1 is estimating σ_1, and S_2 is estimating σ_2.

What if $\sigma_1 = \sigma_2$, but are still unknown?

Pooled T-test

Test Statistic for t-test

$$z = \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

S_1 is estimating σ_1, and S_2 is estimating σ_2.

What if $\sigma_1 = \sigma_2$, but are still unknown?

$\Rightarrow$ They should be estimated by the same estimator S_p.

Pooled T-test

$$
\begin{aligned}
z &= \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} \\
&= \frac{\bar{X} - \bar{Y} - \Delta_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}
\end{aligned}
$$

where S_p is sample standard deviation pooled sample.

Pooled T-test

$$
\begin{aligned}
z &= \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} \\
&= \frac{\bar{X} - \bar{Y} - \Delta_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}
\end{aligned}
$$

where S_p is sample standard deviation pooled sample.

- If assumption of $\sigma_1 = \sigma_2$ are correct, then has slightly higher power than regular t-test.
- It was somewhat popular back then, but it is not a recommended procedure now.

1.4 Paired T-test

If we have random samples from the normal distribution, $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$, but each pair (X_i, Y_i) is a repeated measurements on the same subject, then we can perform paired $t-$test. We first let

$$D_i = X_i - Y_i \qquad \text{for } i = 1, \ldots, n.$$

then perform one-sample $t-$test with test statistic

$$t = \frac{\overline{D} - \Delta_0}{S_D/\sqrt{n}}.$$

This will test the same null hypothesis as $H_0 : \mu_1 - \mu_2 = \Delta_0$ against the alternative of your choice.

H_A	rejection region
upper-tailed	$t > t_{\alpha,n-1}$
lower-tailed	$t < -t_{\alpha,n-1}$
Two-tailed	$t < -t_{\frac{\alpha}{2},n-1}$ or $t > t_{\frac{\alpha}{2},n-1}$

Example: Deer shots

- Study conducted in the Forestry and Wildlife Department at Virginia Tech
- Examined the influence of the drug succinylcholine on the circulation levels of androgens in the blood.
- samples taken from wild deer
- 1st sample: immediately after they had received darts and a capture
- 2nd sample: after 30 min

Test at the 0.05 level of significance whether the androgen concentrations are altered after 30 minutes.

Deer	30_Min_after	Time_of_Injection
1	2.76	7.02
2	5.18	3.10
3	2.68	5.44
4	3.05	3.99
5	4.10	5.21
6	7.05	10.26
7	6.60	13.91
8	4.79	18.53
9	7.39	7.91
10	7.30	4.85
11	11.78	11.10
12	11.58	11.50
13	4.90	3.74
14	22.01	81.03
15	37.38	74.03
16	18.24	31.50
17	13.58	13.10
18	3.70	3.71
19	28.00	44.03
20	67.48	54.03
21	18.94	30.50
22	5.76	6.02
23	5.68	2.80
24	2.68	7.44
25	5.07	3.59
26	4.90	4.21
27	3.90	3.74
28	26.00	94.03
29	67.48	94.03
30	17.04	40.50
31	3.76	5.02
32	5.38	2.80
33	2.68	5.44
34	3.07	3.59

```
35        4.10                  4.21
36        7.25                 10.26
37        6.60                 13.91
38        5.79                 18.53
39        7.39                  7.41
40        7.30                  4.85
```

```
A = c(2.76, 5.18, 2.68, 3.05, 4.10, 7.05, 6.60, 4.79, 7.39, 7.30,
  11.78, 11.58, 4.90, 22.01, 37.38, 18.24, 13.58, 3.70, 28.00,
  67.48, 18.94, 5.76, 5.68, 2.68, 5.07, 4.90, 3.90, 26.00, 67.48,
  17.04, 3.76, 5.38, 2.68, 3.07, 4.10, 7.25, 6.60, 5.79, 7.39, 7.30)

B = c( 7.02, 3.10, 5.44, 3.99, 5.21, 10.26, 13.91, 18.53, 7.91, 4.85, 11.10,
  11.50, 3.74, 81.03, 74.03, 31.50, 13.10, 3.71, 44.03, 54.03, 30.50, 6.02,
   2.80, 7.44, 3.59, 4.21, 3.74, 94.03, 94.03, 40.50, 5.02, 2.80, 5.44, 3.59,
   4.21, 10.26, 13.91, 18.53, 7.41, 4.85)

t.test(A,B)

t.test(A-B)
```

Example: p. 344 Zinc concentration at bottom and top

1.5 Two Sample Inference on Proportion

To test the null hypothesis of $H_0 : p_1 - p_2 = 0$ against alternatives

$$\begin{aligned} H_A: &\quad p_1 - p_2 > 0 \qquad \text{(Upper-tailed alternative)} \\ H_A: &\quad p_1 - p_2 < 0 \qquad \text{(Lower-tailed alternative)} \\ H_A: &\quad p_1 - p_2 \neq 0 \qquad \text{(Two-tailed alternative)}\ , \end{aligned}$$

Use the test statistic

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$\hat{p}_1 = \frac{X}{n_1} \qquad \hat{p}_2 = \frac{Y}{n_2} \qquad \hat{p} = \frac{X+Y}{n_1+n_2}.$$

This is a z-test.

From above characteristics, $100(1-\alpha)\%$ Confidence Interval for $\mu_1 - \mu_2$ can be derived as

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Example: p. 378 Polio vaccine

1954 Salk polio-vaccine double-blind experiment.

Out of 201,229 people who was not vaccinated, 110 got polio.
Out of 200,745 people who was not vaccinated, 33 got polio.

Example: p. 378 Polio vaccine

1954 Salk polio-vaccine double-blind experiment.

Out of 201,229 people who was not vaccinated, 110 got polio.
Out of 200,745 people who was not vaccinated, 33 got polio.

$$\hat{p}_1 = \frac{110}{201,229} = 0.00054664, \quad \hat{p}_2 = \frac{33}{200,745} = 0.00016438$$

Is this a significant difference?

(Polio) two-sample z-test

$H_0 : p_1 - p_2 = 0 \qquad vs. \qquad H_A : p_1 - p_2 > 0$

Perform z-test with

(Polio) two-sample z-test

$H_0 : p_1 - p_2 = 0 \qquad vs. \qquad H_A : p_1 - p_2 > 0$

Perform z-test with

$$\begin{aligned}
\hat{p} &= \frac{33 + 110}{200,745 + 201,229} = 0.00035574. \\
\hat{p}_1 &= 0.00054664, \qquad n_1 = 201,229 \\
\hat{p}_2 &= 0.00016438 \qquad n_2 = 200,745
\end{aligned}$$

(Polio) two-sample z-test

$H_0 : p_1 - p_2 = 0 \qquad vs. \qquad H_A : p_1 - p_2 > 0$
Perform z-test with

$$\begin{aligned} \hat{p} &= \frac{33 + 110}{200,745 + 201,229} = 0.00035574. \\ \hat{p}_1 &= 0.00054664, \qquad n_1 = 201,229 \\ \hat{p}_2 &= 0.00016438 \qquad n_2 = 200,745 \end{aligned}$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = -6.4258$$

Example: Two methods to Check

Suppose that method 1 resulted in 20 unacceptable transistors out of 100 produced; whereas method 2 resulted in 12 unacceptable transistors out of 100 produced. Can we conclude from this, at the 10 percent level of significance, that the two methods are equivalent?

1.6 Two Sample Inference on Variance

Under Case 2,

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(\text{df=(m-1, n-1)})$$

To test the null hypothesis of $H_0 : \sigma_1 = \sigma_2$ against alternatives

$$\begin{aligned}
H_A : &\quad \sigma_1 > \sigma_2 \qquad \text{(Upper-tailed alternative)} \\
H_A : &\quad \sigma_1 < \sigma_2 \qquad \text{(Lower-tailed alternative)} \\
H_A : &\quad \sigma_1 \neq \sigma_2 \qquad \text{(Two-tailed alternative)}\ ,
\end{aligned}$$

we use the test statistic

$$F = \frac{S_1^2}{S_2^2} \sim F(\text{df=(m-1, n-1)}) \qquad \text{if } H_0 \text{ is true}$$

and perform F-test with rejection regions

H_A	rejection region
upper-tailed	$F > \mathcal{F}_{\alpha,m-1,n-1}$
lower-tailed	$F < \mathcal{F}_{1-\alpha,m-1,n-1}$
Two-tailed	$F < \mathcal{F}_{1-\alpha/2,m-1,n-1}$ or $F > \mathcal{F}_{\alpha/2,m-1,n-1}$

CI for Ratio of Variances

$100(1-\alpha)\%$ Confidence Intervals for σ_2^2/σ_1^2

$$\left(\frac{S_2^2}{S_1^2}\mathcal{F}_{1-\frac{\alpha}{2},m-1,n-1},\ \frac{S_2^2}{S_1^2}\mathcal{F}^2_{\frac{\alpha}{2},m-1,n-1}\right) \qquad \text{(two-sided)}$$

$$\left(-\infty,\quad \frac{S_2^2}{S_1^2}\mathcal{F}_{\alpha,m-1,n-1}\right) \qquad \text{(one-sided upper-bound)}$$

$$\left(\frac{S_2^2}{S_1^2}\mathcal{F}_{1-\alpha,m-1,n-1},\ \infty\right) \qquad \text{(one-sided upper-bound)}$$

For $100(1-\alpha)\%$ Confidence Intervals for σ_2/σ_1, take squareroot of above formulas.

Example:

Two types of instruments for measuring the amount of sulfur monoxide in the atmosphere are being compared in an air-pollution experiment. Researchers wish to determine whether the two types of instruments yield measurements having the same variability. The readings in the following table were recorded for the two instruments.

Assuming the populations of measurements to be approximately normally distributed, test the hypothesis that $\sigma_1 = \sigma_2$.

```
Sulfur Monoxide

Instr_A Instr_B
0.86     0.87
0.82     0.74
0.75     0.63
0.61     0.55
0.89     0.76
0.64     0.70
0.81     0.69
0.68     0.57
0.65     0.53
```