Lecture 17

Part 3 Estimation and Hypothesis Test

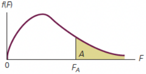
F-distribution

F-distribution

- The null distribution of this F-statistic is an F-distribution with $n_1 - 1$ numerator degrees of freedom and $n_2 - 1$ denominator degrees of freedom.
- The F-distribution is a special continuous distribution:
  - It has two parameters called the numerator degrees of freedom and the denominator degrees of freedom.
  - F-tables give critical values that cut off probability A in the upper tail.
  - There is a different table for each value of A.
  - Rows and columns display the degrees of freedom.

F-distribution

TABLE **6(a)** Critical Values of the *F*-Distribution: $A = .05$



ν_2 \ ν_1	NUMERATOR DEGREES OF FREEDOM 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
DENOMINATOR DEGREES OF FREEDOM 1	161	199	216	225	230	234	237	239	241	242	243	244	245	245	246	246	247	247	248	248
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67	8.67	8.66
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82	5.81	5.80
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58	4.57	4.56
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90	3.88	3.87
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47	3.46	3.44
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17	3.16	3.15
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96	2.95	2.94
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80	2.79	2.77
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67	2.66	2.65
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57	2.56	2.54
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48	2.47	2.46
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41	2.40	2.39
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35	2.34	2.33
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30	2.29	2.28
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26	2.24	2.23
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22	2.20	2.19
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18	2.17	2.16
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.20	2.18	2.17	2.15	2.14	2.12
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.10	2.08	2.07
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.15	2.13	2.11	2.09	2.07	2.05	2.04	2.03
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.02	2.00	1.99
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.09	2.06	2.04	2.02	2.00	1.99	1.97	1.96
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04	2.01	1.99	1.98	1.96	1.95	1.93
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11	2.07	2.04	2.01	1.99	1.96	1.94	1.92	1.91	1.89	1.88
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.97	1.95	1.92	1.90	1.89	1.87	1.85	1.84
45	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.10	2.05	2.01	1.97	1.94	1.92	1.89	1.87	1.86	1.84	1.82	1.81
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.99	1.95	1.92	1.89	1.87	1.85	1.83	1.81	1.80	1.78

Population Variances are Unknown

Testing Equality of Variances

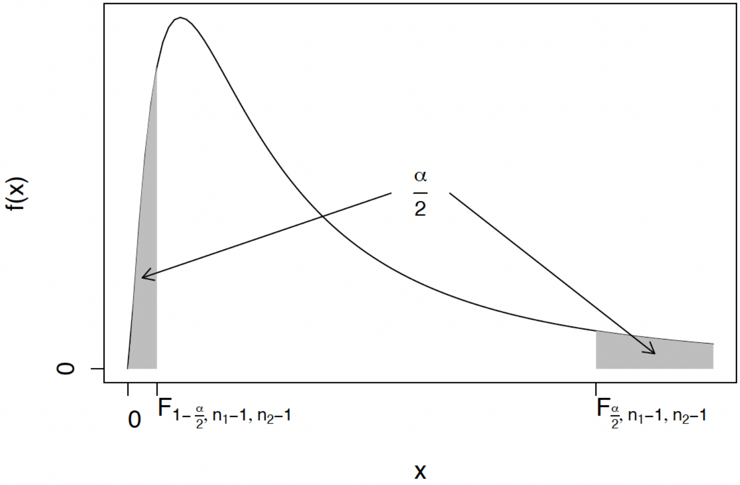
- Decision rule:
  - Compare the F-statistic to the appropriate F-distribution.
  - For a given value of α, reject the null hypothesis if

$$F < F_{1-\frac{\alpha}{2},n_1-1,n_2-1} \text{ or } F > F_{\frac{\alpha}{2},n_1-1,n_2-1}$$

    where $F_{1-\frac{\alpha}{2},n_1-1,n_2-1}$ and $F_{\frac{\alpha}{2},n_1-1,n_2-1}$ are the critical values that cut off $100\left(\frac{\alpha}{2}\right)\%$ in the lower and upper tails, respectively, of the appropriate F-distribution.
  - But the tables don't give us the lower cut-off values.
  - Special trick for doing test...

Population Variances are Unknown

Testing Equality of Variances



Population Variances are Unknown

Testing Equality of Variances

$$F = \frac{s_1^2}{s_2^2}$$

- The labeling of sample 1 and sample 2 is arbitrary, so always put the larger sample variance on top.
- Then our F-statistic will always be larger than 1, and we only have to look up the upper cut-off and reject H_0 at significance level α whenever

$$F > F_{\frac{\alpha}{2}, n_1 - 1, n_2 - 1}.$$

Population Variances are Unknown

Testing $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

- Hypotheses:

$$H_0 : \mu_1 - \mu_2 = D_0$$
$$H_1 : \mu_1 - \mu_2 (\neq, <, >) D_0$$

- Test statistic:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(\overline{X}_1 - \overline{X}_2) - D_0}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Population Variances are Unknown

Testing $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

- We have replaced both population variances in the Z-statistic with s_p^2, the **pooled sample variance**:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- Decision rule:
  - Compare the T-statistic to a t-distribution with $n_1 + n_2 - 2$ degrees of freedom, determine rejection region(s) and decide whether or not to reject H_0.

Population Variances are Unknown

Confidence Interval for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

- A $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$ when the population variances are unknown but equal is given by:

$$(\overline{X}_1 - \overline{X}_2) \pm t_{\frac{\alpha}{2}, n_1+n_2-2} \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Population Variances are Unknown

Testing $\mu_1 - \mu_2$ when $\sigma_1^2 \neq \sigma_2^2$

- What happens when the variances aren't equal?
- The formulae for testing $\mu_1 - \mu_2$ when the variances can't be assumed to be equal are complicated.
- Require the use of computer software to accurately perform the test.
- Therefore they are not considered in this course.

Example

- A company wishes to compare a new type of container with their current type, with respect to the weight of the container.
- Twenty randomly chosen containers of each type are analysed, and some summary statistics are displayed below.

Variable	n	$\overline{X}$	Median	s
Current	20	196.41	198.99	23.20
New	20	183.36	177.49	31.26

Example

Testing Equality of Variances

- Use the output to assess if there is a significant difference between the variances of the weights of the two types of containers.
- Hypotheses:

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

- Test statistic:

$$F = \frac{s_1^2}{s_2^2} = \frac{31.26^2}{23.20^2} = 1.8155$$

Example

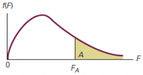
Testing Equality of Variances

- Decision rule:
  - We compare this F-statistic to an F-distribution with $n_1 - 1 = 19$ numerator degrees of freedom and $n_2 - 1 = 19$ denominator degrees of freedom.
  - If we conduct the test at the 5% significance level, we need to look up the F-table for $A = \frac{\alpha}{2} = 0.025$.
  - From the table, the upper cut-off value is 2.53.
- Conclusion:
  - Since $1.8155 < 2.53$, we fail to reject H_0 and we can assume that the variances are equal.

Example

Testing Equality of Variances

TABLE 6(b) Values of the *F*-Distribution: $A = .025$



ν_2 \ ν_1 (NUMERATOR DEGREES OF FREEDOM)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	648	799	864	900	922	937	948	957	963	969	973	977	980	983	985	987	989	990	992	993
2	38.5	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.4
3	17.4	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.4	14.3	14.3	14.3	14.3	14.2	14.2	14.2	14.2	14.2
4	12.2	10.6	10.0	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.79	8.75	8.71	8.68	8.66	8.63	8.61	8.59	8.58	8.56
5	10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.57	6.52	6.49	6.46	6.43	6.40	6.38	6.36	6.34	6.33
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.41	5.37	5.33	5.30	5.27	5.24	5.22	5.20	5.18	5.17
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.71	4.67	4.63	4.60	4.57	4.54	4.52	4.50	4.48	4.47
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.24	4.20	4.16	4.13	4.10	4.08	4.05	4.03	4.02	4.00
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.91	3.87	3.83	3.80	3.77	3.74	3.72	3.70	3.68	3.67
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.66	3.62	3.58	3.55	3.52	3.50	3.47	3.45	3.44	3.42
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.47	3.43	3.39	3.36	3.33	3.30	3.28	3.26	3.24	3.23
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.32	3.28	3.24	3.21	3.18	3.15	3.13	3.11	3.09	3.07
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.95
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.09	3.05	3.01	2.98	2.95	2.92	2.90	2.88	2.86	2.84
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	3.01	2.96	2.92	2.89	2.86	2.84	2.81	2.79	2.77	2.76
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.93	2.89	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.68
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.87	2.82	2.79	2.75	2.72	2.70	2.67	2.65	2.63	2.62
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.81	2.77	2.73	2.70	2.67	2.64	2.62	2.60	2.58	2.56
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.76	2.72	2.68	2.65	2.62	2.59	2.57	2.55	2.53	2.51
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.72	2.68	2.64	2.60	2.57	2.55	2.52	2.50	2.48	2.46
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.65	2.60	2.56	2.53	2.50	2.47	2.45	2.43	2.41	2.39
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.59	2.54	2.50	2.47	2.44	2.41	2.39	2.36	2.35	2.33
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.54	2.49	2.45	2.42	2.39	2.36	2.34	2.31	2.29	2.28
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.49	2.45	2.41	2.37	2.34	2.32	2.29	2.27	2.25	2.23
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.46	2.41	2.37	2.34	2.31	2.28	2.26	2.23	2.21	2.20
35	5.48	4.11	3.52	3.18	2.96	2.80	2.68	2.58	2.50	2.44	2.39	2.34	2.30	2.27	2.23	2.21	2.18	2.16	2.14	2.12
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.13	2.11	2.09	2.07
45	5.38	4.01	3.42	3.09	2.86	2.70	2.58	2.49	2.41	2.35	2.29	2.25	2.21	2.17	2.14	2.11	2.09	2.07	2.04	2.03
50	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38	2.32	2.26	2.22	2.18	2.14	2.11	2.08	2.06	2.03	2.01	1.99

DENOMINATOR DEGREES OF FREEDOM

Example

Testing $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

- Use the output to assess if there is a significant difference between the average weights of the two types of containers.
- Hypotheses:

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0$$

Example

Testing $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

- Test statistic:
  - First calculate the pooled sample variance:

$$
\begin{aligned}
s_p^2 &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \\
&= \frac{19 \times 31.26^2 + 19 \times 23.20^2}{20 + 20 - 2} \\
&= 757.7138
\end{aligned}
$$

Example

Testing $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

- Test statistic continued:
  - Then calculate the T-statistic:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - D_0}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{(183.36 - 196.41) - 0}{\sqrt{757.7138 \times \left(\frac{1}{20} + \frac{1}{20}\right)}}$$

$$= -1.4992$$

Example

Testing $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

- Decision rule:
  - Compare to a t-distribution with $n_1 + n_2 - 2 = 38$ degrees of freedom.
  - Since t-table doesn't list 38 degrees of freedom, use the closest in the table, i.e., 40.
  - For $\alpha = 5\%$, rejection region is $T > 2.021$ or $T < -2.021$ (two-tailed test).
- Conclusion:
  - Since $-2.021 < -1.4992 < 2.021$, there is insufficient evidence to reject the null hypothesis.
  - That is, we conclude that the average weights of the two types of containers are the same.

Paired Samples

Paired Samples

- Paired samples are much easier to deal with than independent samples.
- Let $\{X_{11}, X_{12}, \ldots, X_{1n}\}$ and $\{X_{21}, X_{22}, \ldots, X_{2n}\}$ denote our two paired samples.
- For each pair of observations, calculate the difference $X_{Di} = X_{1i} - X_{2i}$, for $i = 1, \ldots, n$.
- We are now making inferences about $\mu_D = E(X_D)$, the population mean of these differences X_{Di}, and we are in a one-sample situation!

Hypotheses and Test Statistic

- Hypotheses:

$$H_0 : \mu_D = D_0$$
$$H_1 : \mu_D (\neq, <, >) D_0$$

- Test statistic:

$$T = \frac{\overline{X}_D - \mu_D}{\frac{s_D}{\sqrt{n}}} = \frac{\overline{X}_D - D_0}{\frac{s_D}{\sqrt{n}}}$$

- $\overline{X}_D$ and s_D are the sample mean and sample standard deviation, respectively, of the paired differences.

Decision Rule and Conclusion

- Decision rule:
  - We need to compare this T-statistic to a t-distribution with $n-1$ degrees of freedom.
  - Look up the t-table to determine rejection regions.
- Conclusion:
  - If the T-statistic falls in the rejection region, reject H_0. Otherwise, fail to reject H_0.

Confidence Interval

- A $100(1-\alpha)\%$ confidence interval for μ_D, the population mean of the paired differences, is given by:

$$\overline{X}_D \pm t_{\frac{\alpha}{2},n-1}\frac{s_D}{\sqrt{n}}$$

Example

- As a test of how effective antilock brakes are, a car buyer hit the brakes and, using a stopwatch, recorded the number of seconds it took to stop an ABS-equipped car and another identical car without ABS.
- The speeds (mph) when the brakes were applied and the number of seconds each car took to stop are listed on the next slide.
- Can we infer that ABS is better?

Example

Speeds	20	25	30	35	40	45	50	55
ABS	3.6	4.1	4.8	5.3	5.9	6.3	6.7	7.0
Non-ABS	3.4	4.0	5.1	5.5	6.4	6.5	6.9	7.3
Difference	0.2	0.1	-0.3	-0.2	-0.5	-0.2	-0.2	-0.3

Example

- Hypotheses:

$$H_0 : \mu_D = 0$$
$$H_1 : \mu_D < 0$$

- From the differences calculated in the previous table, we can obtain their sample mean and sample standard deviation, $\overline{X}_D = -0.175$ and $s_D = 0.2252$.

Example

- Test statistic:

$$T = \frac{\overline{X}_D - D_0}{\frac{s_D}{\sqrt{n}}} = \frac{-0.175 - 0}{\frac{0.2252}{\sqrt{8}}} = -2.1980$$

- Decision rule and conclusion:
  - Rejection region is $T < t_{0.05,7} = -1.895$.
  - Since $-2.1980 < -1.895$, we reject H_0.
  - That is, there is enough evidence to infer that ABS is better.

Why Product of F-Distribution Critical Values Equals 1

- Let $F_\alpha(df_1, df_2)$ be the upper α critical value of $F(df_1, df_2)$
- We want to prove: $F_\alpha(df_1, df_2) \cdot F_{1-\alpha}(df_2, df_1) = 1$

Why Product of F-Distribution Critical Values Equals 1

$$\text{If } X \sim F(df_1, df_2), \text{ then } \frac{1}{X} \sim F(df_2, df_1)$$

$$P(X > F_\alpha(df_1, df_2)) = \alpha$$

$$P\left(\frac{1}{X} < \frac{1}{F_\alpha(df_1, df_2)}\right) = \alpha$$

$$\frac{1}{F_\alpha(df_1, df_2)} = F_{1-\alpha}(df_2, df_1)$$

$$\therefore F_\alpha(df_1, df_2) \cdot F_{1-\alpha}(df_2, df_1) = 1$$

Test sample proportion

Testing $p_1 - p_2$

- Just as with means, we can test the difference between two population proportions.
- We typically require *independent* samples when dealing with proportions.
- We are making inferences about $p_1 - p_2$, so what is a reasonable estimator of $p_1 - p_2$?
- We can use $\hat{p}_1 - \hat{p}_2$ as an estimator of $p_1 - p_2$, where $\hat{p}_1 = \frac{X_1}{n_1}$ and $\hat{p}_2 = \frac{X_2}{n_2}$.

Testing $p_1 - p_2$

- For large n_1 and n_2, the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is approximately normal with mean and variance given by:

$$E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$$
$$V(\hat{p}_1 - \hat{p}_2) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$$

- As a rule of thumb, this normal approximation is valid provided $n_1 p_1$, $n_1(1-p_1)$, $n_2 p_2$ and $n_2(1-p_2)$ are all larger than 5.

Testing $p_1 - p_2$

- Hypotheses:

$$H_0 : p_1 - p_2 = D_0$$
$$H_1 : p_1 - p_2 (\neq, <, >) D_0$$

  for some $D_0 \neq 0$.

- Test statistic:

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \frac{(\hat{p}_1 - \hat{p}_2) - D_0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

Testing $p_1 - p_2$

- What happens when $D_0 = 0$?
- Hypotheses:

$$H_0 : p_1 - p_2 = 0$$
$$H_1 : p_1 - p_2 (\neq, <, >) 0$$

- We have to change the test statistic slightly.

Testing $p_1 - p_2$

- Test statistic:

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $\hat{p}$ is the **combined proportion**:

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

Confidence Interval for $p_1 - p_2$

- A $100(1-\alpha)\%$ confidence interval for $p_1 - p_2$ is given by:

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

- Note that we can't assume the population proportions are equal when constructing the confidence interval.

Example

- These statistics are calculated from samples taken from two Bernoulli populations:

$$\hat{p}_1 = 0.6, \quad n_1 = 225, \quad \hat{p}_2 = 0.55, \quad n_2 = 225$$

- Calculate the p-value of a test to determine whether there is evidence to infer that the population proportions differ.

Hypotheses and Test Statistic

- Hypotheses:

$$H_0 : p_1 - p_2 = 0$$
$$H_1 : p_1 - p_2 \neq 0$$

- Test statistic:
  - Combined proportion:

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{225 \times 0.6 + 225 \times 0.55}{225 + 225} = 0.575$$

Test Statistic

- Test statistic continued:
  - Z-statistic:

$$
\begin{aligned}
Z &= \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \\
&= \frac{0.6 - 0.55}{\sqrt{0.575(1-0.575)\left(\frac{1}{225} + \frac{1}{225}\right)}} \\
&= 1.0728
\end{aligned}
$$

Decision Rule and Conclusion

- Decision rule
  - The p-value is given by:

$$\begin{aligned} p\text{-value} &= P(Z < -1.07) + P(Z > 1.07) \\ &= 2 \times P(Z < -1.07) \\ &= 0.2846 \end{aligned}$$

- Conclusion:
  - Since 0.2846 is greater than $\alpha = 0.05$ or even $\alpha = 0.1$, we would fail to reject the null hypothesis at both the 5% and 10% significance levels.
  - We conclude the population proportions do not differ.