PUBH 7405-Block 3

LAND ACKNOWLEDGEMENT

The School of Public Health at the University of Minnesota Twin Cities is built within the traditional homelands of the Dakota people. Minnesota comes from the Dakota name for this region, Mni Sóta Makoce, which loosely translates to the land where the waters reflect the skies.

It is important to acknowledge the peoples on whose land we live, learn, and work as we seek to improve and strengthen our relations with our tribal nations. We also acknowledge that words are not enough. We must ensure that our institution provides support, resources, and programs that increase access to all aspects of higher education for our American Indian students, staff, faculty, and community members.

Confidence Intervals and Power

CIs via hypothesis tests

One-sample tests

Recall: $H_0 : E(Y) = \mu_0$ is rejected if:

$$\sqrt{n}|\frac{(\bar{y} - \mu_0)}{s}| \geq t_{1-\alpha/2}(t_{n-1})$$

$H_0 : E(Y) = \mu_0$ is not rejected if:

$$\sqrt{n}|\frac{(\bar{y} - \mu_0)}{s}| \leq t_{1-\alpha/2}(t_{n-1})$$

$$\Rightarrow |\bar{y} - \mu_0| \leq \frac{s}{\sqrt{n}} \times t_{1-\alpha/2}$$

$$\Rightarrow \bar{y} - \frac{s}{\sqrt{n}} t_{1-\alpha/2} \leq \mu_0 \leq \bar{y} + \frac{s}{\sqrt{n}} t_{1-\alpha/2}$$

If μ_0 satisfies the last line, then it is in the "acceptance / non-rejection" region.
If not, it is in the rejection region.

$\Rightarrow$ "plausible" values of μ are in the interval

$$\bar{y} \pm \frac{s}{\sqrt{n}} t_{1-\alpha/2}$$

We call this a $100 \times (1 - \alpha)\%$ confidence interval (CI) for μ.

[contains only those values of μ not rejected by this level-α test]

Main property of CI

Suppose you 1) gather data and 2) compute $100 \times (1 - \alpha)\%$ CI
Also assume $H_0 : E(Y) = \mu_0$ is true.

What is the probability that μ_0 will be in a to-be-sampled (random) interval?

$$P(\mu_0 \text{ in interval } \mid E(Y) = \mu_0)$$

$$= 1 - P(\mu_0 \text{ not in interval } \mid E(Y) = \mu_0)$$

$$= 1 - P(\text{reject } H_0 \mid E(Y) = \mu_0)$$

$$= 1 - P(\text{reject } H_0 \mid H_0 \text{ true})$$

$$= 1 - \alpha$$

$(1 - \alpha)$ is called coverage probability of the interval.

Interpretation:

- pre-experimental probability that the CI will cover the true value
- the large sample fraction of experiments in which the CI covers the true mean.

CI for two sample test

Recall: can construct a 95% CI for $(\mu_B - \mu_A)$ by finding those null hypotheses that would not be rejected at $\alpha = 0.05$ level.

Sampling model:

$$Y_{1A}, \ldots, Y_{n_A A} \sim \text{i.i.d. } \mathcal{N}(\mu_A, \sigma^2)$$

$$Y_{1B}, \ldots, Y_{n_B B} \sim \text{i.i.d. } \mathcal{N}(\mu_B, \sigma^2)$$

Consider whether δ is a reasonable value for the difference in population means.

$$H_0 : \mu_B - \mu_A = \delta$$

$$H_1 : \mu_B - \mu_A \neq \delta$$

Under H_0:

$$\frac{(\bar{Y}_B - \bar{Y}_A) - \delta}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \sim t_{n_A+n_B-2}$$

Thus δ is accepted at level α if observed value

$$\frac{|(\bar{y}_B - \bar{y}_A) - \delta|}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \leq t_c \quad \leftarrow \text{critical value}$$

$$\Rightarrow (\bar{y}_B - \bar{y}_A) - s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}} \leq \delta \leq (\bar{y}_B - \bar{y}_A) + s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$$

$$\text{where} \quad t_c = t_{1-\alpha/2, n_A+n_B-2}$$

From earlier example:

- $\bar{y}_B - \bar{y}_A = 5.93$
- $s_p = 4.72$, $s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}} = 2.72$
- $t_{.975,10} = 2.23$

A 95% confidence interval (CI) for $\mu_B - \mu_A$ is

$$5.93 \pm 2.72 \times 2.23$$

$$5.93 \pm 6.07 = (-0.13, 11.99)$$

Power and Sample Size Determination

Consider

$$H_0 : \mu_A = \mu_B$$

$$H_1 : \mu_A \neq \mu_B$$

Perform a level α test: reject H_0 if

$$|t_{\text{obs}}| \geq t_{1-\alpha/2, n_A+n_B-2}$$

We know if $\alpha = 0.05$:

$$\Rightarrow P(\text{type I error} \mid H_0 \text{ true}) = P(\text{reject } H_0 \mid H_0 \text{ true}) = 0.05$$

What about

$$P(\text{type II error} \mid H_0 \text{ false})?$$

$$= P(\text{do not reject } H_0 \mid H_0 \text{ false})$$

$$= 1 - P(\text{reject } H_0 \mid H_0 \text{ false})$$

Not a well-defined calculation since conditioning argument not clearly specified What is H_0 false?

We need to refer to a specific alternative hypothesis.

$$1 - P(\text{type II error} \mid (\mu_B - \mu_A) = \delta)$$

$$= P(\text{reject } H_0 \mid (\mu_B - \mu_A) = \delta)$$

$$= \text{Power}(\delta)$$

For two sample t-test:

$$\Rightarrow P\left(|t(Y_A, Y_B)| \geq t_{1-\alpha/2, n_A+n_B-2} \mid \delta\right)$$

$\therefore$ we need to know the distribution of our t-statistic under the specific alternative hypothesis.
We need to know the distribution of

$$t(Y_A, Y_B) = \frac{\bar{Y}_B - \bar{Y}_A}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \text{ when } \mu_B - \mu_A = \delta$$

Recall, we know if $\mu_B - \mu_A = \delta$

$$\frac{(\bar{Y}_B - \bar{Y}_A) - \delta}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \sim t_{n_A+n_B-2}$$

but

$$t(Y_A, Y_B) = \frac{(\bar{Y}_B - \bar{Y}_A) - \delta}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} + \frac{\delta}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}$$

move t-statistic away from being centered at zero. $\nearrow$

(amount depends on s_p)

$$\Rightarrow t(Y_A, Y_B) \sim t^*_{n_A+n_B-2}\left(\frac{\delta}{\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}\right)$$

σ is non-centrality parameter $\uparrow$

The Non-central t-distribution

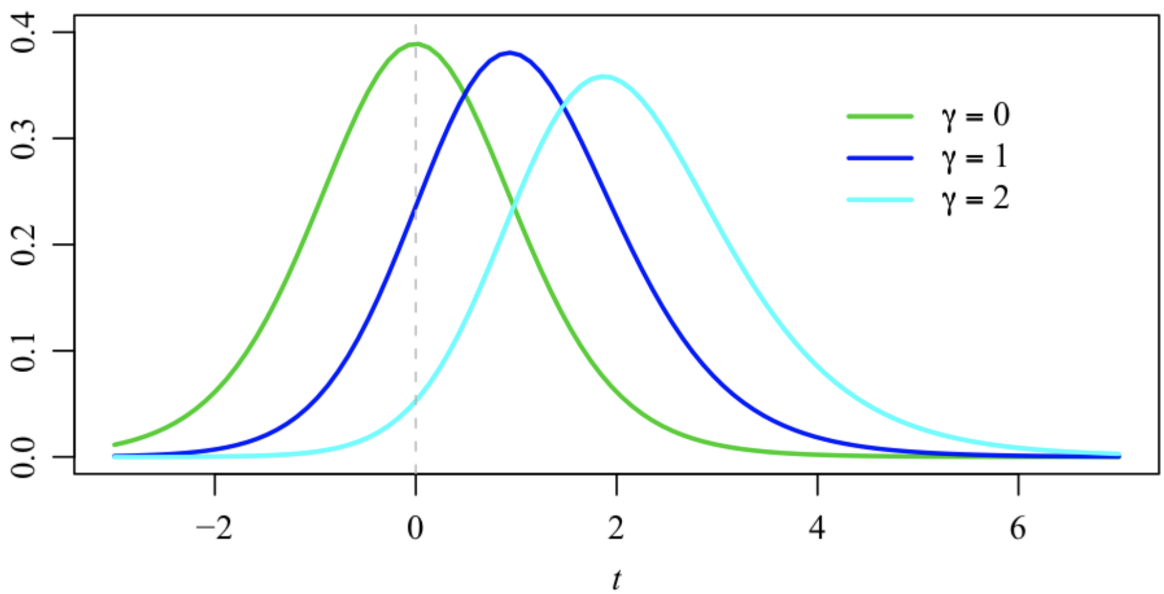
A non-central t-distributed random variable is represented as

$$T = \frac{Z + \delta}{\sqrt{X/\nu}}$$

where δ is a constant, $Z \sim N(0,1)$, $X \sim \chi^2_\nu$

δ is "non-centrality parameter"

Example



A t_{10} distribution and two non-central t_{10}-distributions.

For a non-central t-distribution:

- mean is not zero
- distribution is not symmetric

Recall, if data is normally distributed, with common variance:

$$\bar{Y}_B - \bar{Y}_A \sim \mathcal{N}\left(\delta, \sigma^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)\right)$$

$$\therefore \frac{\bar{Y}_B - \bar{Y}_A}{\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \sim \mathcal{N}\left(\frac{\delta}{\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}, 1\right)$$

Also known as n_A, n_B get large, $s^2 \approx \sigma^2$.

$\therefore$ non-central t-distribution statistic will look approximately normal with the above mean and variance.

Computing the Power of a Test

Recall our level-α testing procedure using two sample t-test:

1. Sample data, compute $t_{\text{obs}} = t(Y_A, Y_B)$
2. Compute p-value: $P\left(|T_{n_A+n_B-2}| > |t_{\text{obs}}|\right)$
3. Reject H_0 if p-value $\leq \alpha$.

We have shown,

$$\begin{gathered} P(\text{reject } H_0 \mid \mu_B - \mu_A = 0) \\ = P(\text{p-value} \leq \alpha \mid \mu_B - \mu_A = 0) \\ = P(|T_{n_A+n_B-2}| \geq t_{1-\alpha/2, n_A+n_B-2}) \\ = \alpha \end{gathered}$$

Now we want:

$$\begin{gathered} P(\text{reject } H_0 \mid \mu_B - \mu_A = \delta) \\ = P\left(|t(Y_A, Y_B)| > t_c \mid \mu_B - \mu_A = \delta\right) \\ = P(|T^*| > t_c) \\ = P(T^* > t_c) + P(T^* < -t_c) \\ = [1 - P(T^* < t_c)] + P(T^* < -t_c) \end{gathered}$$

Where T^* has a non-central t-distribution with n_A+n_B-2 degrees of freedom and non-centrality parameter

$$\Delta? = \frac{\delta}{\sigma\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}}$$

In R:

```
t.crit <- qt( 1 - alpha / 2 , nA + nB - 2 )

t.gamma <- delta / ( sigma * sqrt(1/nA + 1/nB) )

t.power <- 1 - pt( t.crit , nA + nB - 2 , ncp = t.gamma ) +
               pt( -t.crit , nA + nB - 2 , ncp = t.gamma )
```

Approximating Power

Recall for large n_A, n_B:

$$t(Y_A, Y_B) \sim \mathcal{N}(\delta, 1)$$

$$\Rightarrow \text{power given by } P(|X| > t_c) = [1 - P(X < t_c)] + P(X < -t_c)$$

where $X \sim \mathcal{N}(\delta, 1)$

```
t.norm.power <- 1 - pnorm( t.crit, mean = t.gamma ) +
                    pnorm( -t.crit, mean = t.gamma )
```

Sample Size Estimation

How big should the sample size be if we want to reject

$$H_0 : \mu_B - \mu_A = 0 \quad \text{at} \quad \alpha = 0.05$$

with true $\mu_B - \mu_A = 5$ or more?

What we have:

- effect size = 5
- σ^2 unknown, replace with our estimate $s^2 = 22.24 = \hat{\sigma}^2$

Simplify by letting

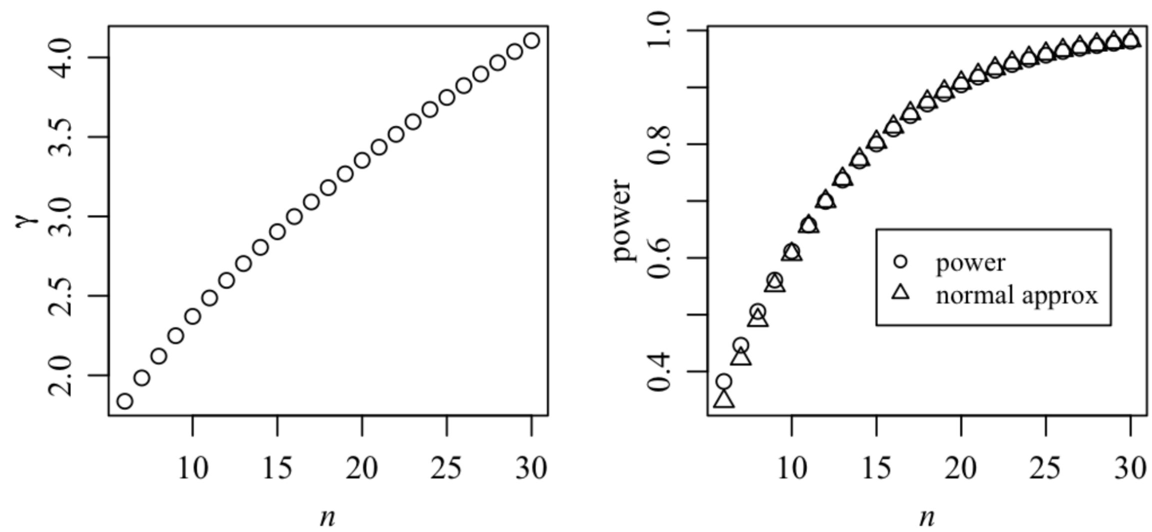
$$n_A = n_B = n$$

$$\Rightarrow \gamma = \frac{\mu_B - \mu_A}{\hat{\sigma}\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \quad (\mu_B - \mu_A \text{ is } \delta)$$

$$= \frac{5}{4.75\sqrt{\frac{2}{n}}} = 0.75\sqrt{n}$$

What is the probability we reject H_0 at level α for a given sample size?

```
delta <- 5 ; s2 <- ( (nA-1)*var(yA) + (nB-1)*var(yB) ) / (nA-1+nB-1)

alpha <- 0.05 ; n <- seq(6,30)

t.crit <- qt(1-alpha/2, 2*n-2)

t.gamma <- delta / sqrt(s2*(1/n+1/n))

t.power <- 1 - pt(t.crit, 2*n-2, ncp=t.gamma) +
           pt(-t.crit, 2*n-2, ncp=t.gamma)

t.normal.power <- 1 - pnorm(t.crit, mean=t.gamma) +
           pnorm(-t.crit, mean=t.gamma)
```



To have 80% power, require

$$n_A = n_B = 15 \quad \text{if true} \quad \mu_B - \mu_A = 5$$

Increasing Power

Recall normal approximation to power, for fixed α, power is a function of

$$\gamma = \frac{\mu_B - \mu_A}{\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}$$

$\therefore$ power is:

- increasing in $|\mu_B - \mu_A|$
- increasing in n_A and n_B
- decreasing in σ^2

Sample size for a Given Precision

To design a study that will pin down the estimate of $\mu_B - \mu_A$ to within $\pm m$ with $1 - \alpha$ confidence, when $n_A = n_B = n$, and when n is large enough so that $t_{2n-2,1-\alpha/2}$ can be approximated by the critical value from $\mathcal{N}(0,1)$

$$[\text{e.g. } z = 1.96 \text{ when } \alpha = 0.05]$$

$$\Rightarrow n = 2\left[\frac{Z\sigma}{m}\right]^2$$

Large-Sample Confidence Interval for p

An approximate CI for the population proportion p is given by:

$$\hat{p} \pm z^* SE \text{ where } SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

z^* is the critical value corresponding to the desired confidence level $Z_{1-\alpha/2}$ where onfidence levels and corresponding z^* values:

Confidence Level	90%	95%	99%
z^*	1.645	1.960	2.576

Remark: The exact SE should be $\sqrt{\frac{p(1-p)}{n}}$, but the unknown p is replaced with the estimate $\hat{p}$. This large-sample CI is not very accurate, meaning the actual confidence level often falls below the nominal level.

Example: Side Effects of Pain Relievers (1)

Arthritis is a painful, chronic inflammation of the joints, so many arthritis patients rely on pain relievers, like Ibuprofen. However, Ibuprofen may induce side effects (like dizziness, muscle cramp, allergy, or even seizure) on some patients.
A study interviewed 440 arthritis patients taking Ibuprofen, and found 23 had experienced side effects. Suppose the 440 patients is a Simple Random Sample (SRS) from the population of arthritis patients taking Ibuprofen.

Find a 90% confidence interval for the population proportion p of arthritis patients who suffer some adverse symptoms.

The sample proportion is $\hat{p} = \frac{23}{440} \approx 0.052$.
The z^* for a 90% CI is 1.645. So a 90%-CI for p is

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \approx 0.052 \pm 1.645 \sqrt{\frac{0.052 \times (1-0.052)}{440}}$$

$$\approx 0.052 \pm 0.017 = (0.035, 0.069)$$

Conclusion: With a 90% confidence, between 3.5% and 6.9% of arthritis patients taking this pain medication experience some adverse symptoms.

Choosing a Sample Size

How large the sample size n need to be to make the margin of error of a CI $\leq m$?

$$\text{margin of error} = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq m \quad \Rightarrow \quad n \geq \left(\frac{z^*}{m}\right)^2 \hat{p}(1-\hat{p})$$

But $\hat{p}$ is UNKNOWN before we get the data. Need to make a guess for p^*. How to choose p^*?

1. Conduct a small pilot study, or use prior studies or knowledge to get a range for possible values of p. Choose the bound that is closer to 0.5. E.g., if possible range of p is [0.1, 0.2], choose $p^* = 0.2$. If the possible range of p is [0.85, 0.95], choose $p^* = 0.85$.

2. The most **conservative** approach is to choose $p^* = 0.5$ since the margin of error is the largest when $\hat{p} = 0.5$.

Example – Sample Size Calculation for a Proportion

A 1993 survey reported that 72.1% of freshmen responding to a national survey were attending the college of their first choice. Suppose that $n = 500$ students responded to the survey.

1. Find a 95% Confidence Interval (CI) for the proportion p of college freshmen attending their first choice college.

2. Suppose that given the CI, we want to conduct a survey which has a margin of error of 1% (i.e., $m = 0.01$) with 95% confidence. How many people should we interview?

The two-sided 95% confidence interval is:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.721 \pm 1.96 \sqrt{\frac{0.721(1-0.721)}{500}}$$

$$= (0.682, 0.760)$$

We have good reason to believe p is in that range. For a sample size calculation, we choose the p closest to 0.5, that is $p = 0.682$.

$$\left(\frac{z^*}{m}\right)^2 p(1-p) = \left(\frac{1.96}{0.01}\right)^2 \times 0.682(1-0.682) \approx 8331.51$$

The required sample size is 8332. We need a much larger sample size than the original study because we want a smaller margin of error.

Remark. Recall that for confidence intervals, we use

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

but for hypothesis testing we use

$$SE = \sqrt{\frac{p_0(1-p_0)}{n}}.$$

Why?

- Recall by CLT when n is large, $\hat{p} \dot{\sim} N(p, \sqrt{\frac{p(1-p)}{n}})$
- When constructing CIs for p, p is unknown, so we estimate $\sqrt{\frac{p(1-p)}{n}}$ by $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
- Under $H_0 : p = p_0$, p is known to be p_0. There is no need to estimate p, and the $\sqrt{\frac{p(1-p)}{n}}$ is simply $\sqrt{\frac{p_0(1-p_0)}{n}}$.

Large Sample Confidence Intervals for $p_1 - p_2$

When n_1 and n_2 are both large,

$$\hat{p}_1 - \hat{p}_2 \dot{\sim} N\left(p_1 - p_2, \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}\right)$$

An approximate $(1-\alpha)100\%$ confidence interval for $p_1 - p_2$ is

$$\text{estimate} \pm z^* \text{SE}$$

where

$$\text{estimate} = \hat{p}_1 - \hat{p}_2, \quad \text{SE} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Use this method only when the number of successes and the number of failures in both samples are at least 10, i.e.,

$$n_1\hat{p}_1, \quad n_1(1-\hat{p}_1), \quad n_2\hat{p}_2, \quad n_2(1-\hat{p}_2) \quad \text{all} \geq 10.$$

Example: Aspirin and Heart Attacks (1)

The Physicians' Health Study was a 5-year randomized study published testing whether regular intake of aspirin reduces mortality from cardiovascular disease[1].

- Participants were male physicians 40-84 years old in 1982 with no prior history of heart attack, stroke, and cancer, no current liver or renal disease, no contraindication of aspirin, no current use of aspirin.
- Every other day, the male physicians participating in the study took either one aspirin tablet or a placebo.
- Response: whether the participant had a heart attack (including fatal or non-fatal) during the 5-year period.

Result:

	Heart Attack?			
Group	Yes	No	Sample Size	
Placebo	189	10845	11034	$\Rightarrow \hat{p}_1 = \frac{189}{11034} \approx 0.0171$
Aspirin	104	10933	11037	$\Rightarrow \hat{p}_2 = \frac{104}{11037} \approx 0.0094$

The z^* for a 99% CI is 2.58, so the 99% CI for $p_1 - p_2$ is

$$\hat{p}_1 - \hat{p}_2 \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$= 0.0171 - 0.0094 \pm 2.58\sqrt{\frac{0.0171(1-0.0171)}{11034} + \frac{0.0094(1-0.0094)}{11037}}$$
$$= 0.0077 \pm 0.0040 = (0.0037, 0.0117)$$

Conclusion:

- As the 99% CI does not contain 0, the incidence rate of heart attack was significantly lower in the aspirin group than in the placebo group (at $\alpha = 0.01$).

[1]Source: Preliminary Report: Findings from the Aspirin Component of the Ongoing Physicians' Health Study. New Engl. J. Med., 318: 262-64, 1988.

- Can we claim that taking aspirin every other day is effective in reducing the chance of heart attack? **Yes, because it was a randomized, double-blind, placebo-controlled experiment.**

Summary: Standard Errors

	One Sample	Two Samples
Mean	$\frac{s}{\sqrt{n}}$	$\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$ if $\sigma_1 \neq \sigma_2$ $s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ if $\sigma_1 = \sigma_2$ where $s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$
proportion (CIs)	$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$
proportion (Tests)	$H_0 : p = p_0$ $\sqrt{\frac{p_0(1-p_0)}{n}}$	$H_0 : p_1 = p_2$ $\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$ where $\hat{p} = \frac{X_1+X_2}{n_1+n_2}$

Use the ones from CIs as part of margin of error calculations.

Sample Size for Comparing Two Proportions

Using the idea of a given precision

$$m = Z_{1-\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

For fixed $n_1 = n_2 = n$, CIs for the independent comparing proportions has maximum width when $p_1 = p_2 = 0.5$.
Why?

Recall

$$\text{var}(\hat{p}_1 - \hat{p}_2) = \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}$$

When $p_1 = p_2 = p$ and $n_1 = n_2 = n$

$$\Rightarrow \frac{\hat{p}(1-\hat{p})}{n} + \frac{\hat{p}(1-\hat{p})}{n}$$

$$= 2\frac{\hat{p}(1-\hat{p})}{n}$$

maximized when $\hat{p} = 0.5$

When this is used, further simplification:

$$\Rightarrow 2 \cdot \frac{0.25}{n} = \frac{1}{2n}$$

Using $\alpha = 0.05\,(Z^* = 1.96)$, worst case margin of error is:

$$m = 1.96\sqrt{\frac{1}{2n}}$$

$$\Rightarrow n = \frac{1.92}{m^2}$$

This gives a formula for the number of subjects in each group n to obtain a given margin of error m.