Hypothesis testing

$$\boxed{X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F_\theta}$$

$$\downarrow$$

H_0 versus H_1 about θ

- Four possibilities in total

		Decision	
		Do not reject H_0	Reject H_0
H_0	True	Correct Decision	Type I Error
	False	Type II Error	Correct Decision

- **Type I error:** reject H_0 when H_0 is true $\rightsquigarrow$ false positive
- **Type II error:** do not reject H_0 when H_0 is false $\rightsquigarrow$ false negative

Example: Chips Ahoy! (revisited)

- $\mu \rightsquigarrow$ the mean number of chips in a bag $H_0 : \mu \leq 1200$ versus $H_1 : \mu > 1200$
- n = 42
- Sample mean = 1261.57
- Sample standard deviation = 117.58
- $\alpha = 0.05$

$$t = \frac{1261.57 - 1200}{117.58/\sqrt{42}} = 3.39 \qquad p = \mathbb{P}(Z \geq 3.39) = 0.00035 \leq 0.05$$

Reject the null hypothesis at level 0.05!

Problem 1. (10 points.) A statistician is worried that a casino is loading a die so that 1 appears more often (than 1/6). To test this, the statistician rolls the die 144 times and get 32 ones. The statistician's null hypothesis is that the die is fair, and the alternative hypothesis is that the 1 shows up more often. The statistician calculates the z-score as

$$\frac{32 - 24}{\sqrt{144 \cdot \frac{1}{6} \cdot \frac{5}{6}}} = 1.79$$

The statistician then obtains a one-sided p-value 0.037 and concludes that there is a 3.7% chance that the die is fair. Do you agree with this conclusion (i.e., do you agree with the interpretation of the p-value)? If not, what would you do differently?

Solution:
I do not agree with the interpretation of the p-value. The correct interpretation should be that if we were to repeatedly sample (rolling the die), we would expect to see a result like this or more extreme about 3.7% of the times if the null hypothesis is true. Given a significant level of $\alpha = 0.05$, I would conclude that there is enough evidence against the null hypothesis. I would reject the null hypothesis.

Testing mean w/ known variance

$$\boxed{X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F,\ \mathbb{E}[X_1] = \mu,\ \mathrm{Var}(X_i) = \sigma^2}$$

$$\mathbb{P}_{H_0}(\text{reject } H_0) = \mathbb{P}_{H_0}(|\bar{X} - \mu_0| \geq C) = \mathbb{P}_{H_0}\left(\frac{|\bar{X} - \mu_0|}{\sigma/\sqrt{n}} \geq C\frac{\sqrt{n}}{\sigma}\right) \approx 2\left(1 - \Phi\left(\frac{C\sqrt{n}}{\sigma}\right)\right)$$

$$\text{Setting } 2\left(1 - \Phi\left(\frac{C\sqrt{n}}{\sigma}\right)\right) = \alpha \Rightarrow \frac{C\sqrt{n}}{\sigma} = z_{\alpha/2} \Rightarrow C = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\boxed{\text{Reject } H_0 \text{ when } |\bar{X} - \mu_0| \geq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}$$

- Among a clinic's patients having high cholesterol levels of at least 240 ml/dl of blood serum, volunteers were recruited to test a new drug designed to reduce blood cholesterol
- A group of 10 volunteers were given the drug for 60 days
- Change of blood cholesterol levels were recorded
- Average change was a decrease of 3.4
- Sample standard deviation is 12.1

Problem 2. (25 points.) An experiment is planned to compare the mean of a control group to the mean of an independent sample of a group given a treatment. Suppose that there are to be $n = 25$ samples in each group. Suppose that the observations are approximately normally distributed and that the standard deviation of a single measurement in either group is $\sigma = 5$.

(a) With a significance level $\alpha = .05$, what is the rejection region of the test of the null hypothesis $H_0 : \mu_X = \mu_Y$ versus the alternative $H_A : \mu_Y > \mu_X$?
(b) What is the power of the test in (a) if $\mu_Y = \mu_X + 1$?
(c) How large should n be so that the test in (a) have a power of at least 0.95 when $\mu_Y = \mu_X + 1$?
(d) What is the rejection region if the alternative is $H_A : \mu_Y \neq \mu_X$? What is the power then if $\mu_Y = \mu_X + 1$?

[Use the normal CDF $\Phi(\cdot)$ in your answer.]

- Signals from the same star are independently received 20 times
- Sample mean = 11.6
- Significance level $\alpha = 0.05$
- Do we reject the H_0?

$$|11.6 - 10| = 1.6 \quad < \quad z_{0.025} \times 4/\sqrt{20} = 1.75$$

Fail to reject the null hypothesis at level 0.05!

- Test the hypothesis that the change in blood cholesterol was purely due to change
- $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$
- Assume normal population
- Significance level $\alpha = 0.05$

$$|\bar{X} - \mu_0| = 3.4 < 8.66 = t_{9,0.025}\frac{S}{\sqrt{10}}$$

Fail to reject the null hypothesis at level 0.05!

Solution:

(a) We have $\mathbf{Var}(\bar{Y} - \bar{X}) = \frac{2\sigma^2}{n} = 2$, and $\mathbb{E}(\bar{Y} - \bar{X}) = 0$ under the null hypothesis. So the rejection region is

$$\frac{\bar{Y} - \bar{X}}{\sqrt{2}} > Z_\alpha$$
$$\Longleftrightarrow \bar{Y} - \bar{X} > \sqrt{2} * 1.64 = 2.32$$

(b)

$$\begin{aligned}\textbf{Power} &= \mathbf{P}\Big(\frac{\bar{Y} - \bar{X}}{\sqrt{\frac{2*25}{n}}} > Z_\alpha | \mu_Y = \mu_X + 1\Big) \\ &= \mathbf{P}\Big(\frac{\bar{Y} - \bar{X} - 1}{\sqrt{\frac{50}{n}}} > Z_\alpha - \frac{1}{\sqrt{\frac{50}{n}}} | \mu_Y = \mu_X + 1\Big) \\ &= \mathbf{P}\Big(\mathbf{N}(0,1) > Z_\alpha - \frac{1}{\sqrt{\frac{50}{n}}}\Big) \\ &= 1 - \Phi\Big(Z_\alpha - \frac{1}{\sqrt{\frac{50}{n}}}\Big) \\ &= \Phi\Big(-Z_\alpha + \frac{1}{\sqrt{\frac{50}{n}}}\Big) \\ &= \Phi\Big(-1.64 + \frac{1}{\sqrt{2}}\Big) \\ &= \Phi(-0.93)\end{aligned}$$

(c)

$$\begin{aligned}\textbf{Power} &= \mathbf{P}\Big(\frac{\bar{Y} - \bar{X}}{\sqrt{\frac{2*25}{n}}} > Z_\alpha | \mu_Y = \mu_X + 1\Big) \\ &= \mathbf{P}\Big(\frac{\bar{Y} - \bar{X} - 1}{\sqrt{\frac{50}{n}}} > Z_\alpha - \frac{1}{\sqrt{\frac{50}{n}}} | \mu_Y = \mu_X + 1\Big) \\ &= \mathbf{P}\Big(\mathbf{N}(0,1) > Z_\alpha - \frac{1}{\sqrt{\frac{50}{n}}}\Big) \\ &= 1 - \Phi\Big(Z_\alpha - \frac{1}{\sqrt{\frac{50}{n}}}\Big) \\ &= \Phi\Big(-Z_\alpha + \frac{1}{\sqrt{\frac{50}{n}}}\Big)\end{aligned}$$

We want $\Phi(-Z_\alpha + \frac{1}{\sqrt{\frac{50}{n}}}) \geq 0.95$, taking the inverse we get

$$\Phi^{-1}\Phi\Big(-Z_\alpha + \frac{1}{\sqrt{\frac{50}{n}}}\Big) = \Phi^{-1}(0.95)$$
$$-1.64 + \frac{1}{\sqrt{\frac{50}{n}}} > 1.64$$
$$\frac{1}{\sqrt{\frac{50}{n}}} > 3.28$$
$$n > 537.9$$

(d) The new rejection region is

$$\frac{\bar{Y} - \bar{X}}{\sqrt{2}} > Z_{\alpha/2} \text{ and } \frac{\bar{Y} - \bar{X}}{\sqrt{2}} < Z_{1-\alpha/2}$$

Problem 5. (20 points.) Life threatening arrhythmias can be predicted from an electrocadiogram by measuring the lengths of QT intervals (the distance from the starts of the Q wave to the starts of the T wave). Suppose we wish to test whether two different calipers, A and B, have the same variability in their measurements. We use these two calipers to measure a set of 8 QT intervals, and the sample variances for the two calipers are 833 and 652, respectively.

(a) Test the hypothesis that the two calipers have different variabilities. Use $\alpha = 0.05$.
(b) Suppose the population standard deviation of caliper A, σ_A is 1.1 times as large as that of caliper B, σ_B ($\sigma_A = 1.1 \cdot \sigma_B$). What is the power of the test in (a)?
[You can use the CDF of the F-distribution $F_{\nu_1,\nu_2}(\cdot)$ in your answer, where ν_1 and ν_2 are the degrees of freedom of the F-distribution.]

Solution:

(a)

$$H_0 : \sigma_A^2 = \sigma_B^2 \quad \text{versus} \quad H_1 : \sigma_A^2 > \sigma_B^2 \,(\text{or } H_1 : \sigma_A^2 \neq \sigma_B^2)$$

or equivalently

$$H_0 : \frac{\sigma_A^2}{\sigma_B^2} = 1 \quad \text{versus} \quad H_1 : \frac{\sigma_A^2}{\sigma_B^2} > 1 \,(\text{or } H_1 : \frac{\sigma_A^2}{\sigma_B^2} \neq 1)$$

Test Statistic:

$$f = \frac{s_A^2}{s_B^2} = \frac{833}{652} \approx 1.2776 < f_{7,7,\alpha} = 3.79 \,(< f_{7,7,\alpha/2})$$

with degrees of freedom $\nu_1 = n_A - 1 = 8 - 1 = 7$ and $\nu_2 = n_B - 1 = 8 - 1 = 7$
Note: The $f_{7,7,\alpha}$ cutoff is from the F-distribution table. The table gives you $\alpha = 0.05$ values. (if you wanted $\alpha = 0.025$ for the two-sided alternative test, realize that failing to reject the H_0 at a higher α level means you also fail to reject for a smaller α).
Conclusion: We don't have statistically significant evidence to reject the H_0 since our test statistic was below the $\alpha/2$ cutoff, which means its p-value is > 0.05. Since we fail to reject the null, we can't say if there exists a difference in the calipers' variances.

(b) Power:

$$\begin{aligned}\mathbb{P}_{\sigma_A = 1.1\cdot\sigma_B}\left(f = \frac{s_A^2}{s_B^2} > f_{7,7,\alpha}\right) &= \mathbb{P}_{\sigma_A = 1.1\cdot\sigma_B}\left(\frac{s_A^2}{s_B^2}\cdot\frac{\sigma_B^2}{\sigma_A^2} > f_{7,7,\alpha}\cdot\frac{\sigma_B^2}{\sigma_A^2}\right) \\ &= \mathbb{P}_{\sigma_A = 1.1\cdot\sigma_B}\left(\frac{s_A^2/\sigma_A^2}{s_B^2/\sigma_B^2} > f_{7,7,\alpha}\cdot\frac{\sigma_B^2}{\sigma_A^2}\right) \\ &= \mathbb{P}_{\sigma_A = 1.1\cdot\sigma_B}\left(\frac{s_A^2/\sigma_A^2}{s_B^2/\sigma_B^2} > \frac{f_{7,7,\alpha}}{(1.1)^2}\right)\end{aligned}$$

Because we assume that

$$\sigma_A = 1.1 \cdot \sigma_B \rightarrow \frac{\sigma_B}{\sigma_A} = \frac{1}{1.1} \rightarrow \frac{\sigma_B^2}{\sigma_A^2} = \frac{1}{(1.1)^2}$$

We also know that

$$\frac{s_A^2/\sigma_A^2}{s_B^2/\sigma_B^2} \sim F_{7,7}$$

so we can say the power for this test is

$$1 - F_{7,7}\left(\frac{f_{7,7,\alpha}}{(1.1)^2}\right) = \mathbf{1 - F_{7,7}\left(\frac{3.79}{(1.1)^2}\right)} \approx 1 - F_{7,7}\,(3.13223)$$

for $F_{7,7}(\cdot)$ as the CDF function for the F-distribution with degrees of freedom $\nu_1 = 7$ and $\nu_2 = 7$.

Testing mean w/ known variance

Normal population or $n \geq 30$

- **Two-sided test**
  - $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0 \rightsquigarrow$ Reject H_0 when $|\bar{X} - \mu_0| \geq z_{\alpha/2}\sigma/\sqrt{n}$
- **One-sided test**
  - $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0 \rightsquigarrow$ Reject H_0 when $\bar{X} - \mu_0 \geq z_\alpha\sigma/\sqrt{n}$
  - $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0 \rightsquigarrow$ Reject H_0 when $\bar{X} - \mu_0 \leq -z_\alpha\sigma/\sqrt{n}$

Test on variance in the normal population

$$\boxed{X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)}$$

$$H_0 : \sigma^2 = \sigma_0^2 \text{ versus } H_1 : \sigma^2 \neq \sigma_0^2$$

Recall that $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ $\Rightarrow$ Under H_0 $\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$

Reject H_0 if $\frac{(n-1)S^2}{\sigma_0^2} \geq \chi^2_{n-1,\alpha/2}$ or $\frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{n-1,1-\alpha/2}$

	Before	After	After - Before
	24.6	10.1	-14.5
	17.0	5.7	-11.3
	16.0	5.6	-10.4
	10.4	3.4	-7.0
	8.2	6.5	-1.7
	7.9	0.7	-7.2
	8.2	6.5	-1.7
	7.9	0.7	-7.2
	5.8	6.1	0.3
	5.4	4.7	-0.7
	5.1	2.0	-3.1
	4.7	2.9	-1.8
Sample mean	10.1	4.58	-5.53
Sample SD	6.06	2.75	4.78

What can you conclude about the effect of captopril? State the assumptions you made in order to arrive at your conclusion.

Solution:
For this problem, we use a significance level = 0.05. First we state the hypotheses.
Null hypothesis: The treatment has no effect or positive effect on the amounts of urinary protein.
Alternative hypothesis: The treatment has a negative effect on the amounts of urinary protein.

Thus, let us compute the test statistic. Since the sample size is small and the variance is unknown, we use t-distribution here.

$$t = \frac{-5.53 - 0}{4.78/\sqrt{12}} = -4.01$$

Note that we are doing a one-sided test here. By the table for t-distribution, we know the rejection rule is

$$t_{11,0.95} = -1.796$$

Since $-4.01 < -1.796$, we reject the null hypothesis that the treatment has no effect or positive effect on the amounts of urinary protein.

Testing mean w/ unknown variance

$$n \geq 30 \rightsquigarrow S \approx \sigma$$

- **Two-sided test**
  - $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0 \rightsquigarrow$ Reject H_0 when $|\bar{X} - \mu_0| \geq z_{\alpha/2}S/\sqrt{n}$
- **One-sided test**
  - $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0 \rightsquigarrow$ Reject H_0 when $\bar{X} - \mu_0 \geq z_\alpha S/\sqrt{n}$
  - $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0 \rightsquigarrow$ Reject H_0 when $\bar{X} - \mu_0 \leq -z_\alpha S/\sqrt{n}$

- Cereal boxes are filled by a manufacturing company using a particular piece of equipment, which filled boxes according to a normal distribution with mean 368 grams and standard deviation 15 grams
- The piece of equipment fails and the manufacturing company orders a replacement.
- 25 cereal boxes produced by the new equipment are selected at random
- The sample standard deviation is computed to be 17.7 grams
- Test whether the standard deviation of the equipment has changed ($\alpha = 0.05$)

Is the average number of chips in a bar greater than 1200?

- $\mu \rightsquigarrow$ the mean number of chips in a bag

$$H_0 : \mu \le 1200 \text{ versus } H_1 : \mu > 1200$$

- Significance level $\alpha = 0.05$
- Rejection rule: reject H_0 when $\bar{X} - \mu_0 \ge z_\alpha S/\sqrt{n}$

$$\bar{X} - \mu_0 = 1261.57 - 1200 = 61.57 > 29.84 = z_{0.05}S/\sqrt{n}$$

Reject the null hypothesis at level 0.05!

$$H_0 : \sigma^2 = 225 \text{ versus } H_1 : \sigma^2 \neq 225$$

$$\text{Reject } H_0 \text{ if } \frac{24 \times S^2}{225} \ge \chi^2_{24,0.025} \text{ or } \frac{24 \times S^2}{225} \le \chi^2_{24,0.975}$$

$$\frac{24S^2}{225} = 33.418 \qquad \chi^2_{24,0.975} = 12.401 \qquad \chi^2_{24,0.025} = 39.364$$

Fail to reject H_0 at level 0.05

New power is

$$\textbf{Power} = \mathbf{P}\Big(\frac{\bar{Y}-\bar{X}}{\sqrt{\frac{2*25}{n}}} > Z_{\alpha/2} | \mu_Y = \mu_X + 1\Big) + \mathbf{P}\Big(\frac{\bar{Y}-\bar{X}}{\sqrt{\frac{2*25}{n}}} < Z_{1-\alpha/2} | \mu_Y = \mu_X + 1\Big)$$
$$= \mathbf{P}\Big(\frac{\bar{Y}-\bar{X}-1}{\sqrt{2}} > Z_{\alpha/2} - \frac{1}{\sqrt{2}} | \mu_Y = \mu_X + 1\Big) + \mathbf{P}\Big(\frac{\bar{Y}-\bar{X}-1}{\sqrt{2}} < Z_{1-\alpha/2} - \frac{1}{\sqrt{2}} | \mu_Y = \mu_X + 1\Big)$$
$$= \mathbf{P}\Big(\mathbf{N}(0,1) > Z_{\alpha/2} - \frac{1}{\sqrt{2}}\Big) + \mathbf{P}\Big(\mathbf{N}(0,1) < Z_{1-\alpha/2} - \frac{1}{\sqrt{2}}\Big)$$
$$= 1 - \Phi\Big(Z_{\alpha/2} - \frac{1}{\sqrt{2}}\Big) + \Phi\Big(Z_{1-\alpha/2} - \frac{1}{\sqrt{2}}\Big)$$

Table 7.6 Level α Tests on σ^2 (Test Statistic: $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$)

Testing Problem	Hypotheses	Reject H_0 if	P-value
Upper One-Sided	$H_0: \sigma^2 \le \sigma_0^2$ vs. $H_1: \sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{n-1,\alpha}$ $\iff$ $s^2 > \frac{\sigma_0^2 \chi^2_{n-1,\alpha}}{n-1}$	$P_U = P(\chi^2_{n-1} \ge \chi^2)$
Lower One-Sided	$H_0: \sigma^2 \ge \sigma_0^2$ vs. $H_1: \sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{n-1,1-\alpha}$ $\iff$ $s^2 < \frac{\sigma_0^2 \chi^2_{n-1,1-\alpha}}{n-1}$	$P_L = P(\chi^2_{n-1} \le \chi^2)$
Two-Sided	$H_0: \sigma^2 = \sigma_0^2$ vs. $H_1: \sigma^2 \neq \sigma_0^2$	$\chi^2 > \chi^2_{n-1,\alpha/2}$ or $\chi^2 < \chi^2_{n-1,1-\alpha/2}$ $\iff$ $s^2 > \frac{\sigma_0^2 \chi^2_{n-1,\alpha/2}}{n-1}$ or $s^2 < \frac{\sigma_0^2 \chi^2_{n-1,1-\alpha/2}}{n-1}$	$2\min\{P_U, P_L = 1 - P_U\}$

Comparing variances

$$\boxed{F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n-1,m-1}}$$

- Testing $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 \neq \sigma_2^2$

$$\frac{S_1^2}{S_2^2} > f_{n-1,m-1,\alpha/2} \text{ or } \frac{S_1^2}{S_2^2} < f_{n-1,m-1,1-\alpha/2}$$

level is shown in the following table.

Educational level	County %	Number of jurors
Elementary	28.4%	1
Secondary	48.5%	10
Some college	11.9%	18
College degree	11.2%	33
Total	100%	62

Test the null hypothesis that the educational level distribution of grand jurors is the same as that of the population (i.e., the county %). Use a 5% significance level.

Solution:
First we state the hypotheses.
Null hypothesis: The educational level distribution of grand jurors is the same as that of the population.
Alternative hypothesis: At least one level of the educational level distribution of grand jurors is different from that of the population.

Firstly, we compute the expected counts of the each education level.
Elementary: $62 \times 28.4\% = 17.61$
Secondary: $62 \times 48.5\% = 30.07$
Some college: $62 \times 11.9\% = 7.38$
College degree: $62 \times 11.2\% = 6.94$

Then, we compute the chi-square test statistic,

$$\chi^2 = \frac{(1-17.61)^2}{17.61} + \frac{(10-30.07)^2}{30.07} + \frac{(18-7.38)^2}{7.38} + \frac{(33-6.94)^2}{6.94} = 142.12$$

The rejection rule is

$$\chi^2_{3,0.05} = 7.81$$

Thus, we reject the null hypothesis.

	Control	Peptic Ulcer
Group A	4219	579
Group O	4578	911

Is there a relationship between blood type and propensity to have peptic ulcer? Use $\alpha = 0.05$.

Solution:

H_0: There is **no relationship** between blood type and propensity to have peptic ulcer.
(i.e. Blood type and the propensity to have peptic ulcer are **independent** of one another.)
v.s.
H_1: There **exists** a relationship between blood type and propensity to have peptic ulcer.
(i.e. Blood type and the propensity to have peptic ulcer are **dependent** on one another.)

- Original Table (with row and column totals)

	Control	Peptic Ulcer	Total
Group A	4219	579	4798
Group O	4578	911	5489
Total	8797	1490	10287

- Independent Table's Expected Count

	Control	Peptic Ulcer
Group A	$\frac{4798\times8797}{10287}$	$\frac{4798\times1490}{10287}$
Group O	$\frac{5489\times8797}{10287}$	$\frac{5489\times1490}{10287}$

$\approx$

	Control	Peptic Ulcer
Group A	4103.043	694.957
Group O	4693.957	795.043

- Test Statistic $\chi^2 = \sum \frac{(\text{Observed}-\text{Expected})^2}{\text{Expected}}$

	Control	Peptic Ulcer
Group A	$\frac{(4219-4103.043)^2}{4103.043}$	$\frac{(579-694.957)^2}{694.957}$
Group O	$\frac{(4578-4693.957)^2}{4693.957}$	$\frac{(911-795.043)^2}{795.043}$

$\approx$

	Control	Peptic Ulcer
Group A	3.277	19.348
Group O	2.865	16.912

- $\chi^2 \approx 3.277 + 19.348 + 2.865 + 16.912 = 42.402 \ge \chi^2_{1,\alpha=0.05} = 3.843$ from the χ^2-distribution table with degrees of freedom $(r-1)(c-1) = (2-1)(2-1) = 1$ (for $r = \#$ rows and $c = \#$ columns).
- Conclusion: We reject the H_0 because our test statistic is above the cutoff (meaning the p-value of this test is $< \alpha = 0.05$). Therefore, there's statistically significant evidence to show there exists *some type of association* between blood type and propensity to have peptic ulcers. (We don't know what the relationship is though.)

	Before	After	After - Before
	24.6	10.1	-14.5
	17.0	5.7	-11.3
	16.0	5.6	-10.4
	10.4	3.4	-7.0
	8.2	6.5	-1.7
	7.9	0.7	-7.2
	8.2	6.5	-1.7
	7.9	0.7	-7.2
	5.8	6.1	0.3
	5.4	4.7	-0.7
	5.1	2.0	-3.1
	4.7	2.9	-1.8
Sample mean	10.1	4.58	-5.53
Sample SD	6.06	2.75	4.78

What can you conclude about the effect of captopril? State the assumptions you made in order to arrive at your conclusion.

Solution:
For this problem, we use a significance level = 0.05. First we state the hypotheses.
Null hypothesis: The treatment has no effect or positive effect on the amounts of urinary protein.
Alternative hypothesis: The treatment has a negative effect on the amounts of urinary protein.

Thus, let us compute the test statistic. Since the sample size is small and the variance is unknown, we use t-distribution here.

$$t = \frac{-5.53 - 0}{4.78/\sqrt{12}} = -4.01$$

Note that we are doing a one-sided test here. By the table for t-distribution, we know the rejection rule is

$$t_{11,0.95} = -1.796$$

Since $-4.01 < -1.796$, we reject the null hypothesis that the treatment has no effect or positive effect on the amounts of urinary protein.

Duality between CI and HT: example

- $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$
- Suppose σ known

At level α, reject H_0 if $\frac{|\bar{X} - \mu_0|}{\sigma/\sqrt{n}} \ge z_{\alpha/2}$

$(1-\alpha)$-CI: $\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$

Reject $H_0 \iff \theta_0 \notin$ CI

Example: coffee and blood flow during exercise

Namdar et al (2006). Caffeine decreases exercise-induced myocardial flow reserve. *Journal of the American College of Cardiology* 47, 405-410.

- Doctors studying healthy men measured myocardial blood flow (MBF) during bicycle exercise after giving the subjects a placebo or a dose of 200 mg of caffeine that was equivalent to drinking two cups of coffee
- 8 subjects $\rightsquigarrow$ each tested twice
- 4 of them randomly selected to receive caffeine in the first test and placebo in the second
- The other 4 received placebo first and caffeine second
- 24-hour gap between the two sets (washout period)

- In the final exam (total points 100):
  - Class 1: mean = 75, SD = 15
  - Class 2: mean = 70, SD = 12
- Is there a statistically significant difference at level $\alpha = 0.05$?

$H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$

$$|\bar{X} - \bar{Y}| = |75 - 70| = 5 \quad > \quad z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = z_{0.025} \times \sqrt{\frac{15^2}{150} + \frac{12^2}{150}} = 3.074$$

We can therefore reject the H_0

Comparison of means with matched pairs

$$D_i \overset{iid}{\sim} \mathcal{N}(\mu_1 - \mu_2, \sigma_D^2)$$

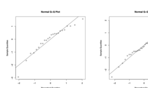
- Sample variance of the differences S_D
- Pivotal statistics

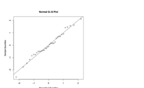
$$\frac{\bar{D} - (\mu_1 - \mu_2)}{S_D/\sqrt{n}} \sim t_{n-1}$$

- Test $H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$

Observed test statistic $t = \frac{\bar{D}}{S_D/\sqrt{n}}$, p-value $= \mathbb{P}(|t_{n-1}| \ge |t|)$

	Perished	Survived
Mean	727.9	738
SD	23.54	19.84
Sample size	24	35





- Do not assume equal variance
- 95% CI for the average arm bone difference between the survival group and the perish group

$$w_1 = 23.54^2/24 = 22.87 \qquad w_2 = 19.85^2/35 = 11.25$$

$$\nu = \frac{(w_1 + w_2)^2}{w_1^2/(n-1) + w_2^2/(m-1)} \approx 43$$

CI=[-1.72,21.92]

Previously in the equal variance case: [-1.27, 21.47]

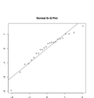
	Married	Single	Widowed or Divorced	Total
Severe	22	16	19	57
Normal	33	29	14	76
Mild	14	9	3	26
Total	69	54	36	159

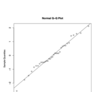
	Married	Single	Widowed or Divorced
Severe	$\frac{57\times69}{159} = 24.7$	$\frac{57\times54}{159} = 19.4$	$\frac{57\times36}{159} = 12.9$
Normal	$\frac{76\times69}{159} = 33.6$	$\frac{76\times54}{159} = 25.8$	$\frac{76\times36}{159} = 17.2$
Mild	$\frac{26\times69}{159} = 11.3$	$\frac{26\times54}{159} = 8.8$	$\frac{26\times36}{159} = 5.9$

$$\chi^2 \approx 6.85 < 9.49 = \chi^2_{4,0.95}$$

Fail to reject H_0

	Perished	Survived
Mean	727.9	738
SD	23.54	19.84
Sample size	24	35





- Assume equal variance
- 95% CI for the average arm bone difference between the survival group and the perish group

$$S^2_{\text{pooled}} = \frac{23 \times 23.54^2 + 34 \times 19.84^2}{24 + 35 - 2} = 458.39$$

$$\left[738 - 727.9 - 2.00 \times \sqrt{458.39}\sqrt{1/24 + 1/35}, 738 - 727.9 + 2.00\sqrt{458.39}\sqrt{1/24 + 1/35}\right] = [-1.26, 21.46]$$

Example: testing uniformity of random digits (cont'd)

Digit	0	1	2	3	4	5	6	7	8	9
Observed Count	12	7	12	7	13	13	7	13	6	10

- p_i: the probability of getting i
- Test $H_0 : p_0 = p_1 = \ldots = p_{10} = \frac{1}{10}$

$$\chi^2 = \frac{(12-10)^2}{10} + \frac{(7-10)^2}{10} + \frac{(12-10)^2}{10} + \frac{(7-10)^2}{10} + \frac{(13-10)^2}{10} + \frac{(3-10)^2}{10} + \frac{(7-10)^2}{10} + \frac{(13-10)^2}{10} + \frac{(6-10)^2}{10} + \frac{(10-10)^2}{10} = 7.8 < 16.92 = \chi^2_{9,0.05}$$

Fail to reject H_0

CI of the population mean

$X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_1, \sigma_1^2)$ $\longleftarrow$ Independent $\longrightarrow$ $Y_1, Y_2, \cdots, Y_m \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_2, \sigma_2^2)$

- Small sample!
- σ_1, σ_2 unknown but $\sigma_1 = \sigma_2 = \sigma$
- Two sided $(1-\alpha)$ CI for $\mu_1 - \mu_2$

$$\left[\bar{X} - \bar{Y} - t_{n+m-2,\alpha/2}S\sqrt{\frac{1}{n} + \frac{1}{m}}, \bar{X} - \bar{Y} + t_{n+m-2,\alpha/2}S\sqrt{\frac{1}{n} + \frac{1}{m}}\right]$$

- Test $H_0 : \mu_1 - \mu_2 = \delta$ versus $\mu_1 - \mu_2 \neq \delta_0$

Reject H_0 when $\left|\bar{X} - \bar{Y} - \delta_0\right| \ge t_{n+m-2,\alpha/2}S\sqrt{\frac{1}{n} + \frac{1}{m}}$

Test of Law of independent assortment

$$\boxed{H_0 : p_1 = 9/16,\ p_2 = 3/16,\ p_3 = 3/16,\ p_4 = 1/16}$$

	Yellow/Round	Green/Round	Yellow/Wrinkled	Green/Wrinkled
Observed	315	108	102	31
Expected	$\frac{9}{16} \times 556 = 312.75$	$\frac{3}{16} \times 556 = 104.25$	$\frac{3}{16} \times 556 = 104.25$	$\frac{1}{16} \times 556 = 34.75$

$$\chi^2 \approx 0.604 < 7.81 = \chi^2_{3,0.95}$$

Fail to reject H_0