Chapter 6

Inferences about One and Two Populations: Categorical Data

We look at statistical inference related to proportions.

6.1 Introduction

We look at examples of categorical populations.

Exercise 6.1 (Categorical Populations)

1. *Butterfly Age Distribution.* The age distribution of butterflies found in Sweden can be thought of as a categorical population.

category	age (in days)	Sweden
1	under 5	6.7%
2	5 to 15	14.1%
3	16 to 65	69.5%
4	over 65	9.7%

   (a) The percentage of butterflies under the age of 5 days in Sweden is $\pi_1 =$ (circle one) **6.7%** / **14.1%** / **69.2%**.

   (b) The percentage of butterflies in category 4 is $\pi_4 =$ (circle one) **9.7%** / **14.1%** / **69.2%**.

   (c) The percentage of butterflies in categories 1 and 2 is $\theta_1 = \pi_1 + \pi_2 =$ (circle one) **9.7%** / **20.8%** / **69.2%**.

   (d) *If* there were one million butterflies in Sweden, the number aged 16 to 65 would be
   (circle one) **67** / **141** / **695** thousand.

(e) If there were *two* million butterflies in Sweden, the number aged 16 to 65 would be
(circle one) **67** / **141** / **1390** thousand.

(f) Since the data here is associated with the *population* of the ages of all butterflies in Sweden, the proportion of butterflies, π_i, $i = 1, 2, 3, 4$, are all examples of
(circle one) **parameters** / **statistics**.

2. *Fathers, Sons and College.* The observed data from a random sample of 85 families in a midwestern city gives the record of college attendance by fathers and their oldest sons is given in the table below.

observed, f	son attended college	son did not attend college	
father attended college	18	12	30
father did not attend college	22	33	55
	40	45	85

(a) **True** / **False** This data can be categorized in the following way.

category	attendance	observed, f	proportion, $\hat{\pi}$
1	both father and son	18	0.21
2	father does, not son	12	0.14
3	not father, son does	22	0.26
4	neither father nor son	33	0.39

(b) The observed number of families in which both the father and son attend college, is
$f_1 =$ (circle one) **12** / **18** / **22**.

(c) The observed number of families in category 3 is
$f_3 =$ (circle one) **12** / **18** / **22**.

(d) The observed proportion of families in category 3 is
$\hat{\pi}_3 =$ (circle one) **0.21** / **0.14** / **0.26**.

(e) The observed proportion of families in category 1 is
$\hat{\pi}_1 =$ (circle one) **0.21** / **0.14** / **0.26**.

(f) The observed proportion of families in which the father attended college is
$\hat{\theta}_1 = \hat{\pi}_1 + \hat{\pi}_2 =$ (circle one) **0.21** / **0.14** / **0.35**.

(g) The observed proportion of families in which the *son* attended college is
$\hat{\theta}_2 =$ (circle one) **0.21** / **0.14** / **0.47**.

(h) Since the data here is associated with a *sample* of families in a midwestern city, the proportion of families in each category, $\hat{\pi}_i$, $i = 1, 2, 3, 4$, are all

examples of
(circle one) **parameters** / **statistics**.
We will use these observed proportions to make inferences on the population proportions.

3. *Greenhouse Sprinkling System.* In a greenhouse, 8% of all plants are assumed to be under–watered. Technical trouble with the greenhouse sprinkling system has raised the concern that the percent under–watered has increased in the past few weeks. Of $n = 60$ plants chosen at random, $\frac{7}{60}$ths ($\frac{7}{60} \approx 0.117$) of them are found to be under–watered.

   (a) **True** / **False** This data can be categorized in the following way.

category	level of watering	observed, f	proportion, $\hat{\pi}$	expected, π
1	under–watered (1)	7	0.117	0.080
2	properly watered (0)	53	0.883	0.920

   (b) The observed number of plants under–watered is
   $f_1 =$ (circle one) **7** / **18** / **22**.

   (c) The observed number of plants in category 2 is
   $f_2 =$ (circle one) **12** / **18** / **53**.

   (d) The *observed* proportion of plants in category 1 is
   $\hat{\pi}_1 =$ (circle one) **0.080** / **0.117** / **0.26**.

   (e) The *expected* proportion of plants in category 1 is
   $\pi_1 =$ (circle one) **0.080** / **0.117** / **0.26**.

   (f) Since the *expected* proportion of plants in category 1 is $\pi_1 = 0.080$, then the *expected number* of plants out of 60 in category 1 is
   $\pi_1 \times 60 = (0.080)(60) =$ (circle one) **1.4** / **4.8** / **6.6**.
   Notice that the observed and expected number of plants in category 1 are not the same.

   (g) **True** / **False** The observed proportion of plants in each category, $\hat{\pi}_i$, $i = 1, 2$, could be used to make inferences (deny or confirm) the expected proportion of plants in each category, π_i, $i = 1, 2$. Since this population involves only one of two possibilities (under–watered (1) or not(0)), it is said to be a *dichotomous* population.

6.2 Proportions From One Population

We look at inferences for proportions from one population.

- *Test and Confidence Interval of Proportion (Two Categories), Large Sample.* For a large (typically, $n \geq 30$) random sample, the test statistic for $H_o : \pi = \pi_o$ is

$$Z = \frac{\hat{\pi} - \pi_0}{\hat{\sigma}_{\hat{\pi}}}, \quad \hat{\sigma}_{\hat{\pi}} = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

  and the (two–sided) confidence interval is given by,

$$\hat{\pi} \pm z(\alpha/2)\hat{\sigma}_{\hat{\pi}}$$

- *Test and Confidence Interval of Proportion (Two Categories), Small Sample.* For a small random sample, the p–values for left, right and two–sided tests for $H_o : \pi = \pi_o$ are given by, respectively,

$$\begin{aligned}
p_- &= f(0|n,\pi_0) + f(1|n,\pi_0) + \cdots + f(r|n,\pi_0) \\
p_+ &= f(n-r|n,\pi_0) + f(n-r+1|n,\pi_0) + \cdots + f(n|n,\pi_0) \\
p &= p_- + p_+, \quad \text{where} \\
& \quad p_- = f(0|n,\pi_0) + f(1|n,\pi_0) + \cdots + f(k|n,\pi_0) \\
& \quad p_+ = f(n-k|n,\pi_0) + f(n-k+1|n,\pi_0) + \cdots + f(n|n,\pi_0) \\
& \quad k = \min(r, n-r)
\end{aligned}$$

  where f is a binomial and the confidence interval is determined using table C.8 (page 824) from the text.

- *Goodness of Fit Test (Two or More Categories), Large Sample.* For a large random sample, the test statistic for $H_o : \pi_1 = \pi_{10}, \ldots, \pi_k = \pi_{k0}$, is

$$\chi^2 = \sum \frac{(f - \hat{f})^2}{\hat{f}}$$

  The expected number of observations, $\hat{f}$, changes according to the number of specified (known) proportion parameters.

From a "big picture" point of view, we look at the statistical inference of one–sample and multiple–sample proportion problems.

		mean μ	variance σ^2	proportion π
	one	large n, 3.7, 3.8, 3.9, 3.10, 4.6 small n, 4.3, 4.6	4.4	**6.2**
sample	two	large n, 3.11 small n, 4.3	4.4	6.3
	multiple	chapters 7, 8, 9	not done	**6.2**, 6.3

For all dichotomous (two category) populations, the statistical inference involves the binomial distribution,

$$f(y) = \frac{n!}{y!(n-y)!} \times \pi^y \times (1-\pi)^{n-y}, \quad y = 0, 1, \ldots, n, \quad 0 \leq \pi \leq 1.$$

Exercise 6.2 (Test and Confidence Interval of Proportion, Large Sample) We look at examples of the hypothesis test and confidence interval for proportion[1] π. For a large (typically, $n \geq 30$) random sample, the test statistic is

$$Z = \frac{\hat{\pi} - \pi_0}{\hat{\sigma}_{\hat{\pi}}}, \quad \hat{\sigma}_{\hat{\pi}} = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

and the (two–sided) confidence interval is given by,

$$\hat{\pi} \pm z(\alpha/2)\hat{\sigma}_{\hat{\pi}}$$

$\boxed{\text{See Lab 8, Test For Proportion, } \pi.}$

1. *Test: Greenhouse Sprinkling System.* In a greenhouse, 8% of all plants are assumed to be under–watered. Technical trouble with the greenhouse sprinkling system has raised the concern that the percent under–watered has increased in the past few weeks. Of $n = 60$ plants chosen at random, $\frac{7}{60}$ths ($\frac{7}{60} \approx 0.117$) of them are found to be under–watered. Does this data support the concern about under–watering at $\alpha = 0.05$?

   (a) *P–Value Versus Level of Significance, Standardized.*

      i. The statement of the test, in this case, is (circle one)

         A. $H_0 : \pi = 0.08$ versus $H_1 : \pi < 0.08$
         B. $H_0 : \pi \leq 0.08$ versus $H_1 : \pi > 0.08$
         C. $H_0 : \pi = 0.08$ versus $H_1 : \pi > 0.08$

      ii. *Test.* The standardized observed proportion is

$$\text{z test statistic } = \frac{\hat{\pi} - \pi}{\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}} = \frac{0.117 - 0.08}{\sqrt{\frac{0.117(0.883)}{60}}}$$

      which equals (circle one) **0.89** / **1.06** / **1.55**. The p–value, the chance the standardized observed proportion is 0.89 or more, guessing the population proportion is 0.08, is given by

$$\text{p–value} = P(Z \geq 0.89)$$

      which equals (circle one) **0.04** / **0.15** / **0.19**.
      (Use 2nd DISTR 2:normalcdf(0.89,E99).)
      The level of significance is given by $\alpha = 0.05$.

[1]The calculator does not calculate $\hat{\sigma}_{\hat{\pi}}$ in the same way as is given in the text.

iii. *Conclusion.* Since the p–value, 0.19, is *greater* than the level of significance, $\alpha = 0.05$, we (circle one) **accept** / **reject** the null guess of 0.08.

(b) *Test Statistic Versus Critical Value, Standardized.*

i. The statement of the test, in this case, is (circle one)

A. $H_0 : \pi = 0.08$ versus $H_1 : \pi < 0.08$

B. $H_0 : \pi \leq 0.08$ versus $H_1 : \pi > 0.08$

C. $H_0 : \pi = 0.08$ versus $H_1 : \pi > 0.08$

ii. *Test.* The standardized observed test statistic is

$$\text{z test statistic} = \frac{\hat{\pi} - \pi}{\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}} = \frac{0.117 - 0.08}{\sqrt{\frac{0.117(0.883)}{60}}} \approx 0.89$$

The critical value is (circle one) **0.04** / **0.15** / **1.65**.
(Use 2nd DISTR 3:invNorm(0.95).)

iii. *Conclusion.* Since the test statistic, 0.89, is smaller than the critical value, 1.65, we (circle one) **accept** / **reject** the null guess of 0.08.

(c) *Population, Sample, Statistic, Parameter.* Match the statistical items with the appropriate parts of this plant example.

terms	plant example
(i) population	**(i)** all (watered or under–watered) plants
(ii) sample	**(ii)** proportion under–watered, of all plants, π
(iii) statistic	**(iii)** 60 (watered or under–watered) plants
(iv) parameter	**(iv)** proportion under–watered, of 60 plants, $\hat{\pi}$

terms	(i)	(ii)	(iii)	(iv)
example				

2. *Confidence Interval: Greenhouse Sprinkling System.* In a random sample of 60 plants, $\frac{7}{60}$ths ($\frac{7}{60} \approx 0.117$) of them are found to be under–watered.

(a) A (two–sided) 95% CI is given by
$\hat{\pi} \pm z(\alpha/2)\hat{\sigma}_{\hat{\pi}} = \hat{\pi} \pm z(\alpha/2)\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = 0.117 \pm z(0.05/2)\sqrt{\frac{0.117(0.883)}{60}} =$
(circle one) $\mathbf{(0.036, 0.198)}$ / $\mathbf{(0.046, 0.208)}$ / $\mathbf{(0.056, 0.218)}$.
(To determine z(0.025), type 2nd DISTR 3:invNorm(0.975) ENTER)

(b) A (two–sided) 99% CI is given by
$\hat{\pi} \pm z(\alpha/2)\hat{\sigma}_{\hat{\pi}} = \hat{\pi} \pm z(\alpha/2)\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = 0.117 \pm z(0.01/2)\sqrt{\frac{0.117(0.883)}{60}} =$
(circle one) $\mathbf{(0.010, 0.224)}$ / $\mathbf{(0.020, 0.234)}$ / $\mathbf{(0.030, 0.244)}$.

Exercise 6.3 (Binomial Test of Proportion, Small Sample) For a small random sample, the p–values for left, right and two–sided tests are given by, respectively[2],

$$\begin{aligned}
p_- &= f(0|n,\pi_0) + f(1|n,\pi_0) + \cdots + f(r|n,\pi_0) \\
p_+ &= f(n-r|n,\pi_0) + f(n-r+1|n,\pi_0) + \cdots + f(n|n,\pi_0) \\
p &= p_- + p_+, \quad \text{where} \\
&\quad p_- = f(0|n,\pi_0) + f(1|n,\pi_0) + \cdots + f(k|n,\pi_0) \\
&\quad p_+ = f(n-k|n,\pi_0) + f(n-k+1|n,\pi_0) + \cdots + f(n|n,\pi_0) \\
&\quad k = \min(r, n-r)
\end{aligned}$$

where the notation "$f(k|n,\pi_0)$" is used to represent the binomial distribution with k successes in n trials where each trial has an observed π_0 chance of succeeding. Also, the confidence interval is determined using table C.8 (page 824) from the text.

1. *Pest Control.* Pesticide X is said to kill 60% of pests. Recent studies seem to show that pests have adapted to this pesticide and so not as many pests are killed. Of $n = 16$ pests subjected to pesticide X, $\frac{10}{16}$ths of them are killed. Does this data support the claim that this pesticide is less effective at $\alpha = 0.05$?

   (a) The statement of the test, in this case, is (circle one)

      i. $H_0 : \pi \leq 0.60$ versus $H_1 : \pi > 0.60$

      ii. $H_0 : \pi \geq 0.60$ versus $H_1 : \pi < 0.60$

      iii. $H_0 : \pi = 0.60$ versus $H_1 : \pi \neq 0.60$

   (b) *Test.* The p–value, the chance the observed proportion is $\frac{10}{16}$ths or less, guessing the population proportion is 0.60, is given by

   $$\text{p–value} = Pr\left(\hat{\pi} \leq \frac{10}{16}\right)$$

   which equals (circle one) **0.47** / **0.53** / **0.67**.
   (Use 2nd DISTR A:binomcdf(16,0.6,10).)
   The level of significance is given by $\alpha = 0.05$.

   (c) *Conclusion.* Since the p–value, 0.67, is greater than the level of significance, $\alpha = 0.05$, we (circle one) **accept** / **reject** the null guess of 0.60.

2. *Seed Germination.* In a sheep cloning experiment, 40% of all fertilized eggs reach the first month of development. Recent improved cloning techniques have raised the possibility that the percentage reaching the first month has increased. Of $n = 11$ fertilized eggs chosen at random, $\frac{5}{11}$ths of them are found to have reached the first month. Does this data support the claim that sheep cloning technology has improved at $\alpha = 0.05$?

[2]We will use our calculators to calculate the various sums of binomial probabilities.

(a) The statement of the test, in this case, is (circle one)

i. $H_0 : \pi \leq 0.40$ versus $H_1 : \pi > 0.40$

ii. $H_0 : \pi \geq 0.40$ versus $H_1 : \pi < 0.40$

iii. $H_0 : \pi = 0.40$ versus $H_1 : \pi \neq 0.40$

(b) *Test.* The p–value, the chance the observed proportion is $\frac{5}{11}$ths or more, guessing the population proportion is 0.40, is given by

$$\text{p–value} = Pr\left(\hat{\pi} \geq \frac{5}{11}\right)$$

which equals (circle one) **0.47** / **0.53** / **0.69**.
(Use 2nd DISTR A:binomcdf(11,0.4,4) and then subtract the result from one: 1− 2nd ANS ENTER.)
The level of significance is given by $\alpha = 0.05$.

(c) *Conclusion.* Since the p–value, 0.47, is greater than the level of significance, $\alpha = 0.05$, we (circle one) **accept** / **reject** the null guess of 0.40.

3. *Potato Rot.* In past surveys of midwest potatoes, 10% were found to have potato rot. Of $n = 15$ potatoes chosen at random, 3 of them are found to be rotten. Does this data support the claim that the percentage of potatoes with potato rot has *changed* (either increased or decreased) at $\alpha = 0.05$?

(a) The statement of the test, in this case, is (circle one)

i. $H_0 : \pi \leq 0.10$ versus $H_1 : \pi > 0.10$

ii. $H_0 : \pi \geq 0.10$ versus $H_1 : \pi < 0.10$

iii. $H_0 : \pi = 0.10$ versus $H_1 : \pi \neq 0.10$

(b) *Test.* The p–value, the chance the observed proportion is either $\frac{3}{15}$ths or less *or* $\frac{15-3}{15} = \frac{12}{15}$ths or more, guessing the population proportion is 0.10, is given by

$$\text{p–value} = Pr\left(\hat{\pi} \leq \frac{3}{15}\right) + Pr\left(\hat{\pi} \geq \frac{12}{15}\right)$$

which equals (circle one) **0.47** / **0.53** / **0.94**.
(Use 2nd DISTR A:binomcdf(15,0.1,3) and then add 1− 2nd DISTR A:binomcdf(15,0.1,11).)
The level of significance is given by $\alpha = 0.05$.

(c) *Conclusion.* Since the p–value, 0.94, is greater than the level of significance, $\alpha = 0.05$, we (circle one) **accept** / **reject** the null guess of 0.10.

4. *Confidence Interval: Potato Rot.* Use table C.8, page 824 of the Rao text.

(a) Of $n = 15$ potatoes chosen at random, 3 of them are found to be rotten. A 95% (two–sided) CI is given by
(circle one) $\mathbf{(0.010, 0.224)}$ / $\mathbf{(0.020, 0.234)}$ / $\mathbf{(0.043, 0.481)}$.

(b) Of $n = 15$ potatoes chosen at random, 3 of them are found to be rotten. A 99% (two–sided) CI is given by
(circle one) **(0.010, 0.224)** / **(0.020, 0.234)** / **(0.0239, 0.5605)**.

(c) Of $n = 11$ potatoes chosen at random, 3 of them are found to be rotten. A 99% (two–sided) CI is given by
(circle one) **(0.010, 0.224)** / **(0.020, 0.234)** / **(0.033, 0.693)**.

(d) Of $n = 11$ potatoes chosen at random, 7 of them are found to be rotten. A 99% (two–sided) CI is given by
(circle one) **(0.010, 0.224)** / **(0.020, 0.234)** / **(0.232, 0.931)**.
Use $n - f = 11 - 7 = 4$ and then subtract the limits from 1.

Exercise 6.4 (Goodness of Fit Test, All Proportions Specified) For a large random sample, the test statistic for $H_o : \pi_1 = \pi_{10}, \ldots, \pi_k = \pi_{k0}$, is

$$\chi^2 = \sum \frac{(f - \hat{f})^2}{\hat{f}}$$

In this case, we assume *all* of the proportion parameters, π_i, $i = 1, 2, \ldots, k$, are specified (known).

See Lab 8, Goodness of Fit Test.

1. *Goodness of Fit Test, Butterfly Age Distribution.* The age distribution of a random sample of 463 butterflies living in Uppsala, a city in Sweden, is compared to the age distribution to *all* of Sweden.

age	Sweden	Uppsala (out of 463)
under 5	6.7%	47
5 to 15	14.1%	75
16 to 65	69.5%	296
over 65	9.7%	45

Test if the age distribution in Uppsala is different from the age distribution for all of Sweden at $\alpha = 0.05$.

(a) *Test Statistic Versus Critical Value, Standardized.*

i. *Statement.* The statement of the test is (circle none, one or more)

A. H_0 : age distribution same
versus H_1 : age distribution different

B. $H_0 : \pi_1 = 0.067, \pi_2 = 0.141, \pi_3 = 0.692, \pi_3 = 0.097$
versus H_1 : age distribution different

C. $H_0 : \pi_1 = 0.067, \pi_2 = 0.141, \pi_3 = 0.692, \pi_3 = 0.097$
versus H_1 : at least one π not equal to null

ii. *Test.* Using,

age	observed, f	expected, $\hat{f}$	$\frac{(f-\hat{f})^2}{\hat{f}}$
under 5	47	31	8.3
5 to 15	75	65	1.5
16 to 65	296	322	2.1
over 65	45	45	0

the observed test statistic is $\chi^2 = \sum \frac{(f-\hat{f})^2}{\hat{f}} = 8.3+1.5+2.1+0 = 11.9$.
The standardized upper critical value at $\alpha = 0.05$ with

$$\text{number of age categories} - 1 = 4 - 1 = 3$$

degrees of freedom, is (circle one) **7.8** / **40.1** / **43.2**
(Use PRGM INVCHI2 ENTER 3 ENTER 0.95 ENTER)

iii. *Conclusion.* Since the test statistic, 11.9, is larger than the critical value, 7.8, we (circle one) **accept** / **reject** the null hypothesis that the age distribution in Uppsala is the same as the age distribution for all of Sweden.

(b) *P–Value Versus Level of Significance, Standardized.*

i. *Statement.* The statement of the test is (circle none, one or more)

A. H_0 : age distribution same
versus H_1 : age distribution different

B. $H_0 : \pi_1 = 0.067, \pi_2 = 0.141, \pi_3 = 0.692, \pi_4 = 0.097$
versus H_1 : age distribution different

C. $H_0 : \pi_1 = 0.067, \pi_2 = 0.141, \pi_3 = 0.692, \pi_4 = 0.097$
versus H_1 : at least one π not equal to null

ii. *Test.* Since the standardized test statistic is $\chi^2 = 11.9$, the p–value is given by

$$\text{p–value} = P(\chi^2 \geq 11.9)$$

which equals (circle one) **0.01** / **0.05** / **0.10**.
(Use 2nd DISTR 2:χ^2cdf(11.7,E99,3).)
The level of significance is 0.05.

iii. *Conclusion.* Since the p–value, 0.01, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that the age distribution in Uppsala is the same as the age distribution for all of Sweden.

2. *Goodness of Fit Test, Flower Colors.* The observed color distribution of a random sample of 1000 roses is compared to the expected (using genetic theory) color distribution of all roses.

color	expected proportion	observed number, f
red	0.925	926
strong–pink	0.025	20
weak–pink	0.025	28
white	0.025	26

Test if the observed color distribution of the random sample of these 1000 roses contradicts the expected (using genetic theory) color distribution of all roses at $\alpha = 0.01$

(a) *Test Statistic Versus Critical Value, Standardized.*

i. *Statement.* The statement of the test is (circle none, one or more)

A. H_0 : color distribution same
versus H_1 : color distribution different

B. $H_0 : \pi_1 = 0.925, \pi_2 = 0.025, \pi_3 = 0.025, \pi_4 = 0.025$
versus H_1 : color distribution different

C. $H_0 : \pi_1 = 0.925, \pi_2 = 0.025, \pi_3 = 0.025, \pi_4 = 0.025$
versus H_1 : at least one π not equal to null

ii. *Test.* Using,

color	f	$\hat{f}$	$\frac{(f-\hat{f})^2}{\hat{f}}$
red	926	925	0.00
strong–pink	20	______	1.00
weak–pink	28	25	0.36
white	26	25	______

the observed test statistic is
$\chi^2 = \sum \frac{(f-\hat{f})^2}{\hat{f}} = 0.00 + 1.00 + 0.36 + 0.04 = 1.40.$
The standardized upper critical value at $\alpha = 0.01$
with $C - 1 = 4 - 1 = 3$ df, is
(circle one) **11.3 / 40.1 / 43.2**
(Use PRGM INVCHI2 ENTER 3 ENTER 0.99 ENTER)

iii. *Conclusion.* Since the test statistic, 1.40, is smaller than the critical value, 11.3, we (circle one) **accept / reject** the null hypothesis that the color distribution of roses agrees with the expected (using genetic theory) color distribution of all roses.

(b) *P–Value Versus Level of Significance, Standardized.*

i. *Statement.* The statement of the test is (circle none, one or more)

A. H_0 : color distribution same
versus H_1 : color distribution different

B. $H_0 : \pi_1 = 0.925, \pi_2 = 0.025, \pi_3 = 0.025, \pi_3 = 0.025$
versus H_1 : color distribution different

C. $H_0 : \pi_1 = 0.925, \pi_2 = 0.025, \pi_3 = 0.025, \pi_3 = 0.025$
versus H_1 : at least one π not equal to null

ii. *Test.* Since the standardized test statistic is $\chi^2 = 1.40$, the p–value is given by

$$\text{p–value} = P(\chi^2 \geq 1.40)$$

which equals (circle one) **0.01** / **0.05** / **0.71**.
(Use 2nd DISTR χ^2cdf(1.4,E99,3).)
The level of significance is 0.01.

iii. *Conclusion.* Since the p–value, 0.71, is larger than the level of significance, 0.01, we (circle one) **accept** / **reject** the null hypothesis that the color distribution of roses agrees with the expected (using genetic theory) color distribution of all roses.

(c) *Population, Sample, Parameter and Statistic.*

terms	flower example
(i) population	**(i)** colors of 1000 roses
(ii) sample	**(ii)** observed χ^2
(iii) statistic	**(iii)** colors of all roses
(iv) parameter	**(iv)** expected χ^2

terms	(i)	(ii)	(iii)	(iv)
flower example				

Exercise 6.5 (Goodness of Fit Test, Some Proportions Specified) Once again, we look at the test statistic for $H_o : \pi_1 = \pi_{10}, \ldots, \pi_k = \pi_{k0}$, for a large sample,

$$\chi^2 = \sum \frac{(f - \hat{f})^2}{\hat{f}}$$

However, this time, the expected number of observations, $\hat{f}$, is altered when only some (possibly, none) of the proportion parameters, π_i, $i = 1, 2, \ldots, k$ are specified (known).

See Lab 8, Goodness of Fit Test, No Proportions Specified.

1. *Flu Symptoms and Aspirin Yet Again, No Proportions Specified.* The *observed* data from a random sample of 354 students from PU/NC in an investigation of the effect of aspirin on reducing flu symptoms is given in the table below.

observed, f	aspirin	no aspirin	subtotals
flu symptoms reduced	120	81	201
flu symptoms not reduced	50	103	153
subtotals	170	184	354

Does the observed data from the random sample of these 354 students contradict what we expect to see if whether or not flu symptoms are reduced is *independent* of whether or not aspirin is taken at $\alpha = 0.01$?

(a) *Preliminary Analysis: Figuring Out The Test Statistic.*

i. **True** / **False** In this study, there are two proportion parameters: the proportion who had reduced flu symptoms and the proportion who took aspirin when they had the flu. *Both* of these parameters are *un*known.

ii. The *expected* number of students whose flu symptoms are reduced when a aspirin is taken (assuming flu symptoms are *independent* of whether or not the aspirin is taken) is calculated in the following way:

$$\text{reduced flu symptoms} \times \text{aspirin} \div 354 = \frac{201 \times 170}{354}$$

which is equal to (circle one) **14.1** / **68** / **96.5**.

iii. The expected number of students whose flu symptoms are reduced when a aspirin is *not* taken, is

$$\text{reduced flu symptoms} \times \textit{not} \text{ given aspirin} \div 354 = \frac{201 \times 184}{354}$$

which is equal to (circle one) **12.5** / **73.6** / **104.5**.

iv. The expected number of students whose flu symptoms are *not* reduced when a aspirin is taken, is

$$\text{no reduced flu symptoms} \times \text{aspirin} \div 354$$

which is equal to (circle one) $\mathbf{0.6 \times 170}$ / $\mathbf{\frac{153 \times 170}{354}}$ / $\mathbf{201 \times 0.5}$.

v. The expected number of students whose flu symptoms are *not* reduced when a aspirin is *not* taken, is

$$\text{no reduced flu symptoms} \times \text{not given aspirin} \div 354$$

which is equal to (circle one) $\mathbf{\frac{153 \times 170}{354}}$ / $\mathbf{\frac{153 \times 184}{354}}$ / $\mathbf{0.6 \times 184}$.

vi. **True** / **False** The expected table, created from the observed data, is given below.

expected, $\hat{f}$	aspirin	no aspirin	subtotals
flu symptoms reduced	96.5	104.5	201
flu symptoms not reduced	73.5	79.5	153
subtotals	170	184	354

vii. Consequently, to determine the test statistic,

flu study	observed, f	expected, $\hat{f}$	$\frac{(f-\hat{f})^2}{\hat{f}}$
flu reduced, aspirin given	120	96.5	5.7
flu reduced, aspirin not given	81	104.5	5.2
flu not reduced, aspirin given	50	73.5	7.5
flu not reduced, aspirin not given	103	79.5	6.9

and so $\sum \frac{(f-\hat{f})^2}{\hat{f}}$ = (circle one) **3.84** / **25.3** / **109.2**

(b) *Test Statistic Versus Critical Value, Standardized.*

i. *Statement.* The statement of the test is (circle none, one or more)

A. H_0 : flu symptoms independent of aspirin
versus H_1 : flu symptoms *not* independent of aspirin

B. H_0 : flu symptoms independent of aspirin
versus H_1 : flu symptoms dependent on aspirin

C. H_0 : flu symptoms *not* independent of aspirin
versus H_1 : flu symptoms independent of aspirin

ii. *Test.*
The test statistic is $\sum \frac{(f-\hat{f})^2}{\hat{f}} = 25.3$.
The standardized upper critical value at $\alpha = 0.01$ with

$$\text{number of categories} - 1 - \text{number of unknown parameters}$$
$$= C - 1 - q = 4 - 1 - 2 = 1 \text{ df}$$

is (circle one) **3.84** / **6.63** / **11.3**
(Use PRGM INVCHI2 ENTER 1 ENTER 0.99 ENTER)

iii. *Conclusion.* Since the test statistic, 25.3, is larger than the critical value, 6.63, we (circle one) **accept** / **reject** the null hypothesis that whether or not there is a reduction of flu symptoms is *in*dependent on whether or not the aspirin is taken.

2. *Fathers, Sons and College, No Proportions Specified.* The observed data from a random sample of 85 families in a midwestern city gives the record of college attendance by fathers and their oldest sons is given in the table below.

observed, f	son attended college	son did not attend college	
father attended college	18	12	30
father did not attend college	22	33	55
	40	45	85

Test whether or not a son attends college is *dependent* on whether or not the father attends college at $\alpha = 0.05$.

(a) *Test Statistic Versus Critical Value, Standardized.*

i. *Statement.* The statement of the test is (circle none, one or more)
   A. H_0 : son attends independent of father attending
      versus H_1 : son attends *not* independent of father attending
   B. H_0 : son attends independent of father attending
      versus H_1 : son attends dependent on father attending
   C. H_0 : son attends *not* independent of father attending
      versus H_1 : son attends independent of father attending

ii. Using,

attendance	observed, f	expected, $\hat{f}$	$\frac{(f-\hat{f})^2}{\hat{f}}$
both father and son	18	14.1	1.08
not father, son does	22	25.9	0.59
father does, not son	15	15.9	0.96
neither father nor son	33	29.1	0.52

the observed test statistic is $\sum \frac{(f-\hat{f})^2}{\hat{f}} = 1.08+0.59+0.96+0.52 = 3.15$.
The standardized upper critical value at $\alpha = 0.05$ with

$$\text{number of categories } -1- \text{ number of unknown parameters}$$
$$= C - 1 - q = 4 - 1 - 2 = 1 \text{ df}$$

is (circle one) **3.84** / **40.1** / **43.2**
(Use PRGM INVCHI2 ENTER 1 ENTER 0.95 ENTER)

iii. *Conclusion.* Since the test statistic, 3.15, is smaller than the critical value, 3.84, we (circle one) **accept** / **reject** the null hypothesis that whether or not a son attends college is *independent* on whether or not the father attends college

(b) *P–Value Versus Level of Significance, Standardized.*

i. *Statement.* The statement of the test is (circle none, one or more)
   A. H_0 : son attends independent of father attending
      versus H_1 : son attends *not* independent of father attending
   B. H_0 : son attends independent of father attending
      versus H_1 : son attends dependent on father attending
   C. H_0 : son attends *not* independent of father attending
      versus H_1 : son attends independent of father attending

ii. *Test.* Since the standardized test statistic is $\chi^2 = 3.15$, with 1 df, the p–value is given by

$$\text{p–value} = P(\chi^2 \geq 3.15)$$

which equals (circle one) **0.01** / **0.08** / **0.10**.
(Use 2nd DISTR 2:χ^2cdf(3.15,E99,1).)
The level of significance is 0.05.

iii. *Conclusion.* Since the p–value, 0.08, is larger than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that whether or not a son attends college is *independent* on whether or not the father attends college

3. *Flu Symptoms and Aspirin Again, One Proportion Specified.* The *observed* data from a random sample of 354 students from PU/NC in an investigation of the effect of aspirin on reducing flu symptoms is given in the table below.

observed, f	aspirin	no aspirin	subtotals
flu symptoms reduced	120	22	142
flu symptoms not reduced	50	162	212
subtotals	170	184	354

We know (or, at least, we assume we know) that 40% of *all* students at PU/NC experience reduced flu symptoms for reasons other than taking aspirin (and so 60% do not experience reduced flu symptoms). Does the observed data from the random sample of these 354 students contradict what we expect to see if whether or not flu symptoms are reduced is *independent* of whether or not aspirin is taken at $\alpha = 0.01$?

(a) *Preliminary Analysis: Figuring Out The Test Statistic.*

i. **True** / **False** In this study, there are two proportion parameters: the proportion who had reduced flu symptoms and the proportion who took aspirin when they had the flu. One of these parameters is known (the proportion who had reduced flu symptoms) and one is *un*known (the proportion who took aspirin when they had the flu).

ii. The *expected* number of students whose flu symptoms are reduced when a aspirin is taken (assuming flu symptoms are *independent* of whether or not the aspirin is taken) is calculated in the following way:

$$\begin{aligned} & \text{assumed proportion reduced flu symptoms} \times \text{number taking aspirin} \\ = \; & 0.4 \times 170 \end{aligned}$$

which is equal to (circle one) **14.1** / **68** / **100.5**.

iii. The expected number of students whose flu symptoms are reduced when a aspirin is *not* taken, is

$$\begin{aligned} & \text{assumed proportion reduced flu symptoms} \times \text{number } \textit{not} \text{ taking aspirin} \\ = \; & 0.4 \times 184 \end{aligned}$$

which is equal to (circle one) **12.5** / **73.6** / **153**.

iv. The expected number of students whose flu symptoms are *not* reduced when a aspirin is taken, is

assumed proportion no reduced flu symptoms×number taking aspirin

which is equal to (circle one) $\mathbf{0.6 \times 170}$ / $\mathbf{\frac{153\times170}{354}}$ / $\mathbf{201 \times 0.5}$.

v. The expected number of students whose flu symptoms are *not* reduced when a aspirin is *not* taken, is

assumed proportion no reduced flu symptoms×number not taking aspirin

which is equal to (circle one) $\mathbf{\frac{153\times170}{354}}$ / $\mathbf{\frac{153\times184}{354}}$ / $\mathbf{0.6 \times 184}$.

vi. **True** / **False** The expected table, created from the observed data, is given below.

expected, $\hat{f}$	aspirin	no aspirin	subtotals
flu symptoms reduced	68	73.6	141.6
flu symptoms not reduced	102	110.4	212.4
subtotals	170	184	354

knowing that 40% have reduced flu symptoms and 60% do not have reduced flu symptoms.

vii. Consequently, to determine the test statistic,

flu study	observed, f	expected, $\hat{f}$	$\frac{(f-\hat{f})^2}{\hat{f}}$
flu reduced, aspirin given	120	68	39.8
flu reduced, aspirin not given	22	73.6	36.2
flu not reduced, aspirin given	50	102	26.5
flu not reduced, aspirin not given	162	110.4	24.1

and so $\sum \frac{(f-\hat{f})^2}{\hat{f}}$ = (circle one) **3.84** / **6.63** / **31.6**

(b) *Test Statistic Versus Critical Value, Standardized.*

i. *Statement.* The statement of the test is (circle none, one or more)

A. H_0 : flu symptoms independent of aspirin
versus H_1 : flu symptoms *not* independent of aspirin

B. H_0 : flu symptoms independent of aspirin
versus H_1 : flu symptoms dependent on aspirin

C. H_0 : flu symptoms *not* independent of aspirin
versus H_1 : flu symptoms independent of aspirin

ii. *Test.*
The test statistic is $\sum \frac{(f-\hat{f})^2}{\hat{f}} = 31.6$.
The standardized upper critical value at $\alpha = 0.01$ with

number of categories $-1-$ number of unknown parameters
$= C - 1 - q = 4 - 1 - 1 = 2$ df

is (circle one) **3.84** / **6.63** / **9.21**
(Use PRGM INVCHI2 ENTER 2 ENTER 0.99 ENTER)

iii. *Conclusion.* Since the test statistic, 31.6, is larger than the critical value, 9.21, we (circle one) **accept** / **reject** the null hypothesis that whether or not there is a reduction of flu symptoms is *in*dependent on whether or not the aspirin is taken.

6.3 Comparing Proportions in Two Populations

We will look at *comparing* either two or more proportions in either two independent or two paired populations.

- *Test and Confidence Interval of Difference in Proportions (Independent Populations), Large Sample.* For a large (typically, $n \geq 30$) random sample, the test statistic for $H_o : \theta = \pi_1 - \pi_2 = \theta_0$ is

$$z = \frac{\hat{\theta} - \theta_0}{\hat{\sigma}_{\hat{\theta}}}, \quad \hat{\sigma}_{\hat{\theta}} = \sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}$$

  and the (two–sided) confidence interval is given by,

$$\hat{\theta} \pm z(\alpha/2)\hat{\sigma}_{\hat{\theta}}$$

- *Fisher's Test: Test of Difference in Proportions (Independent Populations), Small Sample.* The p–values for left, right and two–sided tests for $H_o : \pi_1 - \pi_2$ are given by, respectively,

$$\begin{aligned}
p_- &= f(y|0, n_1, n) + f(y|1, n_1, n) + \cdots + f(y|y, n_1, n) \\
p_+ &= f(n_1|n_1 - y, n_1, n) + f(n_1|n_1 - y + 1, n_1, n) + \cdots + f(n_1|n_1, n_1, n) \\
p &= p_- + p_+, \quad \text{where} \\
& \quad p_- = f(y|0, n_1, n) + f(y|1, n_1, n) + \cdots + f(y|y, n_1, n) \\
& \quad p_+ = f(n_1|n_1 - y, n_1, n) + f(n_1|n_1 - y + 1, n_1, n) + \cdots + f(n_1|n_1, n_1, n)
\end{aligned}$$

  where f is a hypergeometric.

- *McNemar's Test: Test and Confidence Interval of Difference in Proportions (Dependent Populations), Large Sample.* For a large (typically, $n \geq 30$) random sample, the test statistic for $H_o : \theta = \pi_1 - \pi_2 = \theta_0$ is

$$z = \frac{\hat{\theta} - \theta_0}{\hat{\sigma}_{\hat{\theta}}}, \quad \hat{\theta} = \frac{B - C}{N}, \quad \hat{\sigma}_{\hat{\theta}} = \frac{\sqrt{B + C}}{N}$$

and the (two–sided) confidence interval is given by,

$$\hat{\theta} \pm z(\alpha/2)\hat{\sigma}_{\hat{\theta}}$$

- *Goodness of Fit Test (Two or More Categories) For Two Populations, Large Sample.* For a large random sample, the test statistic for $H_o : \pi_{11} = \pi_{21}, \ldots, \pi_{1k} = \pi_{2k}$, is

$$\chi^2 = \sum \frac{(f - \hat{f})^2}{\hat{f}}$$

  where it is assumed there are no known proportion parameters.

From a "big picture" point of view, we look at the statistical inference of two–sample or multiple–sample proportion problems.

		mean μ	variance σ^2	proportion π
	one	large n, 3.7, 3.8, 3.9, 3.10, 4.6 small n, 4.3, 4.6	4.4	6.2
sample	two	large n, 3.11 small n, 4.3	4.4	**6.3**
	multiple	chapters 7, 8, 9	not done	6.2, **6.3**

Exercise 6.6 (Tests and Confidence Intervals for Specified Differences in Proportions From Independent Populations, Large Sample[3]) We look at examples of the hypothesis test and confidence interval for the difference in proportions, $\theta = \pi_1 - \pi_2$. For a large (typically, $n \geq 30$) random sample, the test statistic is

$$z = \frac{\hat{\theta} - \theta_0}{\hat{\sigma}_{\hat{\theta}}}, \quad \hat{\sigma}_{\hat{\theta}} = \sqrt{\frac{\hat{\pi}_1(1 - \hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1 - \hat{\pi}_2)}{n_2}}$$

and the (two–sided) confidence interval is given by,

$$\hat{\theta} \pm z(\alpha/2)\hat{\sigma}_{\hat{\theta}}$$

1. *Left Sided Test: Male Urban and Rural Pigeons.* A comparison of the number of male pigeons in urban and rural areas was undertaken, with the following results, based on large random samples.

	urban (1)	rural (2)	total
male pigeons	358	6786	7144
female pigeons	49	577	626
total	407	7363	7770

[3]Can*not* use TESTS 2–Prop ZTest... since the calculator uses *pooled* $\hat{\pi}$, not separate $\hat{\pi}_1$, $\hat{\pi}_2$.

Does this data support the claim there is a smaller proportion of male pigeons in the urban than in rural areas at $\alpha = 0.05$?

(a) *Test Statistic Versus Critical Value, Standardized.*

i. *Statement.* If the proportion of male pigeons in the urban, π_1, is *less* than the proportion of male pigeons in rural life, π_2, then the statement of the test is (circle one)

A. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$
B. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$
C. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

ii. *Test.* The standardized test statistic of

$$\hat{\pi}_1 - \hat{\pi}_2 = \frac{358}{407} - \frac{6786}{7363} = 0.8796 - 0.9216 =$$

(circle one) **−0.042** / **−0.076** / **−0.123**,
with standard error given by,

$$\hat{\sigma}_{\hat{\pi}} = \sqrt{\frac{0.8796(1 - 0.8796)}{407} + \frac{0.9216(1 - 0.9216)}{7363}} =$$

(circle one) **0.0164** / **0.0174** / **0.0184**,
is given by,

$$z \text{ test statistic } = \frac{(\hat{\pi}_1 - \hat{\pi}_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}} = \frac{-0.042 - 0}{0.0164} =$$

(circle one) **−1.23** / **−2.56** / **−4.56**.
The standardized *lower* critical value at $\alpha = 0.05$ is
(circle one) **−1.18** / **−1.65** / **−2.33**
(Use 2nd DISTR 3:invNorm(0.05))

iii. *Conclusion.* Since the test statistic, -2.56, is smaller than the critical value, -1.65, we (circle one) **accept** / **reject** the null hypothesis that $\pi_1 - \pi_2 = 0$.

(b) *P–Value Versus Level of Significance, Standardized.*

i. *Statement.* The statement of the test is (circle one)

A. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$
B. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$
C. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

ii. *Test.* Since the standardized test statistic is $z = -2.56$, the p–value is given by

$$\text{p–value} = P(Z \leq -2.56)$$

which equals (circle one) **0.001** / **0.005** / **0.011**.
(Use 2nd DISTR 2:normalcdf($-$E99,-2.56).)
The level of significance is 0.05.

iii. *Conclusion.* Since the p–value, 0.005, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that $\pi_1 - \pi_2 = 0$.

2. *Two–Sided Test: Male Urban and Rural Pigeons.* A comparison of the number of male pigeons in urban and rural areas was undertaken, with the following results[4], based on large random samples.

	urban (1)	rural (2)	total
male pigeons	358	6786	7144
female pigeons	49	577	626
total	407	7363	7770

Does this data support the claim there is a *different* proportion of male pigeons in the urban than in rural areas at $\alpha = 0.05$?

(a) *Test Statistic Versus Critical Value, Standardized.*

i. *Statement.* If the proportion of male pigeons in the urban, π_1, is *different than* than the proportion of male pigeons in rural life, π_2, then the statement of the test is (circle one)

A. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$

B. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$

C. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

ii. *Test.* The standardized test statistic is given by $z = -2.56$.
The standardized *lower* critical value at $\alpha/2 = 0.025$ is
(circle one) **−1.18** / **−1.65** / **−1.96**
(Use 2nd DISTR 3:invNorm(0.025))
The standardized *upper* critical value at $\alpha/2 = 0.025$ is
(circle one) **1.18** / **1.65** / **1.96**
(Use 2nd DISTR 3:invNorm(0.975))

iii. *Conclusion.* Since the test statistic, -2.56, is *outside* the lower critical value, -1.96, and upper critical value, 1.96, we (circle one) **accept** / **reject** the null hypothesis that $\pi_1 - \pi_2 = 0$.

(b) *P–Value Versus Level of Significance, Standardized.*

i. *Statement.* The statement of the test is (circle one)

A. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$

[4] This the *same* data as above!

B. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$

C. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

ii. *Test.* Since the standardized test statistic is $z = -2.56$, the p–value is given by

$$\text{p–value} = P(Z \leq -2.56) + P(Z \geq 2.56)$$

which equals (circle one) **0.002** / **0.005** / **0.010**.
(Use 2nd DISTR 2:normalcdf(−E99,-2.56) and multiply by two.)
The level of significance is 0.05.

iii. *Conclusion.* Since the p–value, 0.010, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that $\pi_1 - \pi_2 = 0$.

3. *Confidence Intervals: Male Urban and Rural Pigeons.* A comparison of the number of male pigeons in urban and rural areas was undertaken, with the following results, based on large random samples.

	urban (1)	rural (2)	total
male pigeons	358	6786	7144
female pigeons	49	577	626
total	407	7363	7770

(a) A (two–sided) 95% CI is given by
$\hat{\theta} \pm z(\alpha/2)\hat{\sigma}_{\hat{\theta}} = -0.042 \pm z(0.05/2)(0.0164) =$ (circle one)
(−0.0741, −0.0099)
(−0.0841, −0.0199)
(−0.0941, −0.0299).
(To determine z(0.025), type 2nd DISTR 3:invNorm(0.975) ENTER)

(b) A (two–sided) 99% CI is given by
$\hat{\theta} \pm z(\alpha/2)\hat{\sigma}_{\hat{\theta}} = -0.042 \pm z(0.01/2)(0.0164) =$ (circle one)
(−0.0842, −0.0002)
(−0.0942, −0.0102)
(−0.1042, −0.0202).

Exercise 6.7 (Fisher's Exact Test for Specified Differences in Proportions From Independent Populations, Small Sample) We now look at examples of Fisher's hypothesis test for comparing two proportions π_1 and π_2, when the sample size is small. The p–values for left, right and two–sided tests for $H_o : \pi_1 - \pi_2$ are given by, respectively,

$$p_- = f(y|0, n_1, n) + f(y|1, n_1, n) + \cdots + f(y|y, n_1, n)$$

$$\begin{aligned} p_+ &= f(n_1|n_1-y,n_1,n)+f(n_1|n_1-y+1,n_1,n)+\cdots+f(n_1|n_1,n_1,n)\\ p &= p_- + p_+, \quad \text{where}\\ &\quad p_- = f(y|0,n_1,n)+f(y|1,n_1,n)+\cdots+f(y|y,n_1,n)\\ &\quad p_+ = f(n_1|n_1-y,n_1,n)+f(n_1|n_1-y+1,n_1,n)+\cdots+f(n_1|n_1,n_1,n) \end{aligned}$$

where f is a hypergeometric[5] and where y, n_1 and n are defined in the following table,

	(1)	(2)	total
(1)	y	b	n_1
(2)	c	d	n_2
total	m_1	m_2	n

See Lab 8: Hypergeometric Distribution

1. *Working With The Hypergeometric.* A comparison of the number of male and female pigeons in urban and rural areas was undertaken, with the following results, based on small random samples.

	urban (1)	rural (2)	total
male pigeons	y	b	6
female pigeons	c	d	8
total	10	4	14

   (a) It is possible that $\{y,b,c,d\} = \{4,2,6,2\}$ because, for example, the sum of the number of urban and rural male pigeons, $y+b = 4+2$, equals the required total number of male pigeons, 6. Other possible values of $\{y,b,c,d\}$ include (circle none, one or more)

      i. $\{5,1,5,3\}$
      ii. $\{3,3,7,1\}$
      iii. $\{2,4,8,0\}$
      iv. $\{6,0,4,4\}$

   (b) **True** / **False** The value y cannot be any larger than six (6), the minimum of the male pigeon row total, 6, and the urban pigeon column total, 10.

   (c) **True** / **False** The value y cannot be any smaller than two (2), the maximum of zero (0) or[6] sum of the male pigeon total plus the urban pigeon column total minus the total number of pigeons, $6+10-14=2$.

   (d) The probability $\Pr(Y=4)$; in other words, $\{y,b,c,d\} = \{4,2,6,2\}$ is equal to $\frac{10!4!6!8!}{14!4!2!6!2!}$ = (circle one) **0.42** / **0.45** / **0.47**
   (Type PRGM HYPPDF ENTER 14 6 10 4 ENTER)

[5] We have not discussed the hypergeometric; this is the first time we have come across it. We will use our calculators to determine probabilities required for the hypergeometric distribution.

[6] ... or $6-4=2$ or $10-8=2$.

(e) The probability $\Pr(Y = 3)$; in other words, $\{y, b, c, d\} = \{3, 3, 7, 1\}$ is equal to $\frac{10!4!6!8!}{14!3!3!7!1!}$ = (circle one) **0.16** / **0.18** / **0.22**
(Type PRGM HYPPDF ENTER 14 6 10 3 ENTER)

(f) The probability $\Pr(Y \leq 3)$; in other words, $\{y, b, c, d\} = \{3, 3, 7, 1\}$ or $\{y, b, c, d\} = \{2, 4, 8, 0\}$ is equal to
$\frac{10!4!6!8!}{14!3!3!7!1!} + \frac{10!4!6!8!}{14!2!4!8!0!}$ = (circle one) **0.16** / **0.17** / **0.22**
(Type PRGM HYPCDF ENTER 14 6 10 3 ENTER; notice HYPCDF, rather than HYPPDF, is used here!)

(g) **True** / **False** The probability $\Pr(Y = 4)$ is *equal* to the probability $\Pr(B = 2)$ since $\{y, b, c, d\} = \{4, 2, 6, 2\}$ in both cases.

2. *Left Sided Test: Male Urban and Rural Pigeons.* A comparison of the number of male pigeons in urban and rural areas was undertaken, with the following results, based on small random samples.

	urban (1)	rural (2)	total
male pigeons	4	2	6
female pigeons	6	2	8
total	10	4	14

Does this data support the claim there is a smaller proportion of male pigeons in the urban than in rural areas at $\alpha = 0.05$?

(a) *Statement.* The statement of the test is (circle one)

i. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$
ii. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$
iii. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

(b) *Test.* The p–value, the chance the observed number of male pigeons in the urban areas is smaller than or equal to four (4), is given by

$$\text{p–value} = Pr(Y \leq 4)$$

which equals (circle one) **0.47** / **0.53** / **0.59**.
(Use PRGM HYPCDF ENTER 14 6 10 4 ENTER)
The level of significance is given by $\alpha = 0.05$.

(c) *Conclusion.* Since the p–value, 0.59, is greater than the level of significance, $\alpha = 0.05$, we (circle one) **accept** / **reject** the null guess of $\pi_1 - \pi_2 = 0$.

3. *Two–Sided Test: Male Urban and Rural Pigeons.* A comparison of the number of male pigeons in urban and rural areas was undertaken, with the following results, based on small random samples.

	urban (1)	rural (2)	total
male pigeons	6	2	8
female pigeons	1	8	9
total	7	10	17

Does this data support the claim there is either a smaller or larger proportion of male pigeons in the urban than in rural areas at $\alpha = 0.10$?

(a) *Statement.* The statement of the test is (circle one)

i. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$

ii. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$

iii. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

(b) *Test.* The p–value, the chance the observed number of male pigeons in the urban areas is larger than or equal to six (6), or smaller than or equal to two ($8 - 6 = 2$), is given by[7]

$$\text{p–value} = Pr(Y \leq 2) + Pr(Y \geq 6)$$

which equals (circle one) **0.05** / **0.07** / **0.23**.
(Use PRGM HYPCDF ENTER 17 8 7 2 ENTER (which is 0.2178) and add to this 1− PRGM HYPCDF ENTER 17 8 7 5 ENTER (which is 0.0133))
The level of significance is given by $\alpha = 0.10$.

(c) *Conclusion.* Since the p–value, 0.23, is greater than the level of significance, $\alpha = 0.10$, we (circle one) **accept** / **reject** the null guess of $\pi_1 - \pi_2 = 0$.

Exercise 6.8 (McNemar Test and Confidence Interval for Specified Differences in Proportions From Dependent Populations, Large Sample) We look at examples of the hypothesis test and confidence interval for the *paired* difference in proportions $\theta = \pi_1 - \pi_2$. For a large (typically, $n \geq 30$) random sample, the test statistic for $H_o : \theta = \pi_1 - \pi_2 = \theta_0$ is

$$z = \frac{\hat{\theta} - \theta_0}{\hat{\sigma}_{\hat{\theta}}}, \quad \hat{\theta} = \frac{B - C}{N}, \quad \hat{\sigma}_{\hat{\theta}} = \frac{\sqrt{B + C}}{N}$$

and the (two–sided) confidence interval is given by,

$$\hat{\theta} \pm z(\alpha/2)\hat{\sigma}_{\hat{\theta}}$$

and where B, C and N are defined in the following table,

[7] If we had compared males to females, it would have been $7 - 6 = 1$.

	(1)	(2)	total
(1)	A	B	N_1
(2)	C	D	N_2
total	M_1	M_2	N

1. *Right–Sided Test: Flu Symptoms.* The observed data from a random sample of patients in an investigation of the effect of a new drug on reducing flu symptoms is given in the table below. All of the 354 patients were given *both* the drug and also the placebo (no drug). Both treatments lasted 12 weeks. Each patient was in the study for a total of 24 weeks.

drug ↓ placebo →	reduced	not reduced	subtotals
flu symptoms reduced	120	81	201
flu symptoms not reduced	50	103	153
subtotals	170	184	354

Does this data support the claim there is a larger proportion of patients with reduced flu symptoms who took the drug than patients with reduced flu symptoms who took the placebo at $\alpha = 0.05$?

(a) *Test Statistic Versus Critical Value, Standardized.*

i. *Statement.* If the proportion of patients with reduced flu symptoms who took the drug, π_1, is *greater* than the proportion of patients with reduced flu symptoms who took the placebo, π_2, then the statement of the test is (circle one)

A. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$
B. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$
C. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

ii. *Test.* The standardized test statistic of
$\hat{\theta} = \frac{81-50}{354}$ = (circle one) **0.042** / **0.076** / **0.088**,
with standard error given by,
$\hat{\sigma}_{\hat{\pi}} = \frac{\sqrt{81+50}}{354}$ = (circle one) **0.032** / **0.045** / **0.078**,
is given by,

$$Z = \frac{\hat{\theta} - \theta_0}{\hat{\sigma}_{\hat{\theta}}} = \frac{0.088 - 0}{0.032}$$

(circle one) **1.23** / **2.75** / **4.56**.
The standardized upper critical value at $\alpha = 0.05$ is
(circle one) **1.18** / **1.65** / **2.33**
(Use 2nd DISTR 3:invNorm(0.95))

iii. *Conclusion.* Since the test statistic, 2.75, is larger than the critical value, 1.65, we (circle one) **accept** / **reject** the null hypothesis that $\pi_1 - \pi_2 = 0$.

(b) *P–Value Versus Level of Significance, Standardized.*

i. *Statement.* The statement of the test is (circle one)

A. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$

B. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$

C. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

ii. *Test.* Since the standardized test statistic is $z = 2.75$, the p–value is given by

$$\text{p–value} = P(Z \geq 2.75)$$

which equals (circle one) **0.001** / **0.003** / **0.011**.
(Use 2nd DISTR 2:normalcdf(2.75, E99).)
The level of significance is 0.05.

iii. *Conclusion.* Since the p–value, 0.003, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that $\pi_1 - \pi_2 = 0$.

2. *Right–Sided Test: Flu Symptoms.* The observed data from a random sample of patients in an investigation of the effect of a new drug on reducing flu symptoms is given in the table below. All of the 354 patients were given *both* the drug and also the placebo (no drug). Both treatments lasted 12 weeks. Each patient was in the study for a total of 24 weeks.

drug ↓ placebo →	reduced	not reduced	subtotals
flu symptoms reduced	130	71	201
flu symptoms not reduced	40	113	153
subtotals	170	184	354

Does this data support the claim there is a larger proportion of patients with reduced flu symptoms who took the drug than patients with reduced flu symptoms who took the placebo at $\alpha = 0.05$?

(a) *Test Statistic Versus Critical Value, Standardized.*

i. *Statement.* If the proportion of patients with reduced flu symptoms who took the drug, π_1, is *greater* than the proportion of patients with reduced flu symptoms who took the placebo, π_2, then the statement of the test is (circle one)

A. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$

B. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$

C. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

ii. *Test.* The standardized test statistic of
$\hat{\theta} = \frac{71-40}{354}$ = (circle one) **0.042** / **0.076** / **0.088**,

with standard error given by,
$\hat{\sigma}_{\hat{\pi}} = \frac{\sqrt{71+40}}{354}$ = (circle one) **0.030** / **0.045** / **0.078**,
is given by,

$$Z = \frac{\hat{\theta} - \theta_0}{\hat{\sigma}_{\hat{\theta}}} = \frac{0.088 - 0}{0.030}$$

(circle one) **1.23** / **2.93** / **4.56**.
The standardized upper critical value at $\alpha = 0.05$ is
(circle one) **1.18** / **1.65** / **2.33**
(Use 2nd DISTR 3:invNorm(0.95))

iii. *Conclusion.* Since the test statistic, 2.93, is larger than the critical value, 1.65, we (circle one) **accept** / **reject** the null hypothesis that $\pi_1 - \pi_2 = 0$.

(b) *P–Value Versus Level of Significance, Standardized.*

i. *Statement.* The statement of the test is (circle one)
A. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$
B. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 < 0$
C. $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 > 0$

ii. *Test.* Since the standardized test statistic is $z = 2.93$, the p–value is given by

$$\text{p–value} = P(Z \geq 2.93)$$

which equals (circle one) **0.001** / **0.002** / **0.011**.
(Use 2nd DISTR 2:normalcdf(2.93, E99).)
The level of significance is 0.05.

iii. *Conclusion.* Since the p–value, 0.001, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that $\pi_1 - \pi_2 = 0$.

3. *Confidence Intervals: Flu Symptoms.* The observed data from a random sample of patients in an investigation of the effect of a new drug on reducing flu symptoms is given in the table below. All of the 354 patients were given *both* the drug and also the placebo (no drug). Both treatments lasted 12 weeks. Each patient was in the study for a total of 24 weeks.

drug ↓ placebo →	reduced	not reduced	subtotals
flu symptoms reduced	130	71	201
flu symptoms not reduced	40	113	153
subtotals	170	184	354

(a) A (two–sided) 95% CI is given by $\hat{\theta} \pm z(\alpha/2)\hat{\sigma}_{\hat{\theta}} = 0.088 \pm z(0.05/2)(0.032) =$ (circle one) $\mathbf{(0.025, 0.151)}$ / $\mathbf{(0.035, 0.161)}$ / $\mathbf{(0.045, 0.171)}$.
(To determine z(0.025), type 2nd DISTR 3:invNorm(0.975) ENTER)

(b) A (two–sided) 99% CI is given by $\hat{\theta} \pm z(\alpha/2)\hat{\sigma}_{\hat{\theta}} = 0.088 \pm z(0.01/2)(0.032) =$ (circle one) $\mathbf{(0.006, 0.170)}$ / $\mathbf{(0.005, 0.171)}$ / $\mathbf{(0.008, 0.172)}$.

4. *Notation.* Consider the following 2×2 table for the McNemar test.

population 1 ↓ pop 2 →	(1)	(2)	subtotals
(1)	A	B	N_1
(2)	C	D	N_2
subtotals	M_1	M_2	N

(a) $\hat{\theta} =$ (circle one) $\boldsymbol{\frac{B+C}{N}}$ / $\boldsymbol{\frac{B-C}{N}}$ / $\boldsymbol{\frac{C-D}{N}}$

(b) $\hat{\sigma}_{\hat{\theta}} =$ (circle one) $\boldsymbol{\frac{\sqrt{B+C}}{N}}$ / $\boldsymbol{\frac{\sqrt{B-C}}{N}}$ / $\boldsymbol{\frac{\sqrt{C-D}}{N}}$

Exercise 6.9 (Tests and Confidence Intervals for Overall Differences in Proportions From Independent Populations, Large Sample: Plant Growth) The *observed* data from a random sample of 1000 plants in an investigation of the effect of nutritional level on plant growth is given in the table below.

	nutritional level →	poor	adequate	excellent	row totals
plant growth	below average	70	95	35	200
	average	130	450	30	610
	above average	90	30	70	190
	column totals	290	575	135	1000

1. *Preliminary Work.* Complete the following table.

plant study	f	$\hat{f}$	$\frac{(f-\hat{f})^2}{\hat{f}}$
below plant, poor nutrition	70	58.0	2.48
average plant, poor nutrition	130	$\frac{(290)(610)}{1000} =$ ______	12.43
above plant, poor nutrition	90	55.1	22.11
below plant, adequate nutrition	95	115	3.48
average plant, adequate nutrition	450	350.8	28.08
above plant, adequate nutrition	30	$\frac{(575)(190)}{1000} =$ ______	57.49
below plant, excellent nutrition	35	27	2.37
average plant, excellent nutrition	30	82.4	33.28
above plant, excellent nutrition	70	25.7	76.68

2. *Below Average Versus Average Growth: Test Statistic Versus Critical Value.* Test whether the proportion of below average plant growth is the same as the average plant growth at the poor, adequate and excellent nutritional levels at $\alpha = 0.05$.

(a) *Statement.* The statement of the test is (circle one)

i. $H_0 : \pi_{11} = \pi_{21},\ \pi_{12} = \pi_{22},\ \pi_{13} = \pi_{23}$
versus H_1 : at least one of the proportions not equal to one another

ii. $H_0 : \pi_{11} = \pi_{21},\ \pi_{12} = \pi_{22},\ \pi_{13} = \pi_{23}$
versus $H_1 : \pi_{11} \neq \pi_{21},\ \pi_{12} = \pi_{22},\ \pi_{13} = \pi_{23}$

iii. $H_0 : \pi_{11} = \pi_{21},\ \pi_{12} \neq \pi_{22},\ \pi_{13} = \pi_{23}$
versus $H_1 : \pi_{11} \neq \pi_{21},\ \pi_{12} \neq \pi_{22},\ \pi_{13} \neq \pi_{23}$

(b) *Test.* the observed test statistic is
$\sum \frac{(f-\hat{f})^2}{\hat{f}} = 2.48 + 12.43 + 3.48 + 28.08 + 2.37 + 33.28$,
which equals[8] (circle one) **82.12** / **98.33** / **138.46**.
The standardized upper critical value at $\alpha = 0.05$ with

$$\text{number of nutritional levels } -1 = 3 - 1 = 2 \text{ df}$$

is (circle one) **3.84** / **5.99** / **9.49**
(Use PRGM INVCHI2 ENTER 2 ENTER 0.95 ENTER)

(c) *Conclusion.* Since the test statistic, 82.12, is larger than the critical value, 5.99, we (circle one) **accept** / **reject** the null hypothesis that the proportions are all the same.

3. *Below Average Versus Above Average Growth: P–Value Versus Level of Significance.* Test whether the proportion of average plant growth is the same as the excellent average plant growth at the poor, adequate and excellent nutritional levels at $\alpha = 0.05$.

(a) *Statement.* The statement of the test is (circle none, one or more)

i. $H_0 : \pi_{11} = \pi_{31},\ \pi_{12} = \pi_{32},\ \pi_{13} = \pi_{33}$
versus H_1 : at least one of the proportions not equal to one another

ii. $H_0 : \pi_{11} = \pi_{31},\ \pi_{12} = \pi_{32},\ \pi_{13} = \pi_{33}$
versus $H_1 : \pi_{11} \neq \pi_{31},\ \pi_{12} = \pi_{32},\ \pi_{13} = \pi_{33}$

iii. $H_0 : \pi_{11} = \pi_{31},\ \pi_{12} = \pi_{32},\ \pi_{13} = \pi_{33}$
versus $H_1 : \pi_{11} \neq \pi_{31},\ \pi_{12} \neq \pi_{32},\ \pi_{13} \neq \pi_{33}$

(b) *Test.* Since the standardized test statistic[9] is
$\chi^2 = 2.48 + 22.11 + 3.48 + 57.49 + 2.37 + 76.68 =$
(circle one) **82.12** / **98.33** / **164.61**.
with 2 df, the p–value is given by

$$\text{p–value} = P(\chi^2 \geq 164.61)$$

[8]Notice that only the below average and average proportions, but not above average proportions are being used in the summation here!

[9]Notice that only the below average and above average proportions, but not average proportions are being used in the summation here!

which equals (circle one) **0.00** / **0.08** / **0.10**.
(Use 2nd DISTR 2:χ^2cdf(164.61,E99,2).)
The level of significance is 0.05.

(c) *Conclusion.* Since the p–value, 0.00, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that the proportions are all the same.

4. *Notation.* The *observed* data from a random sample of 1000 plants in an investigation of the effect of nutritional level on plant growth is given in the table below.

	nutritional level →	poor	adequate	excellent	row totals
plant	below average	70	95	35	200
growth	average	130	450	30	610
	above average	90	30	70	190
	column totals	290	575	135	1000

(a) If $f_{1+} = 200$, then $f_{2+} =$ (circle one) **290** / **575** / **610**.

(b) If $f_{+1} = 290$, then $f_{+2} =$ (circle one) **290** / **575** / **610**.

(c) $f_{++} =$ (circle one) **290** / **575** / **1000**.

(d) If $\sum_{i=1}^{2} f_{i1} = 290$ then $\sum_{i=1}^{2} f_{i2} =$ (circle one) **290** / **575** / **610**.

(e) **True** / **False** $f_{+1} = \sum_{i=1}^{2} f_{i1}$.

(f) If $\hat{f}_{12} = \frac{f_{1+}f_{+2}}{f_{++}}$, then $\hat{f}_{13} =$ (circle one) $\boldsymbol{\frac{f_{3+}f_{+1}}{f_{++}}}$ / $\boldsymbol{\frac{f_{1+}f_{+3}}{f_{++}}}$ / $\boldsymbol{\frac{f_{1+}f_{+3}}{f_{1+}}}$.