Chapter 7

Inference for proportions

Exercises

Contents

7 Inference for proportions

Exercises **1**
7.1 Passing proportions 3
7.2 Outdoor lighting 9
7.3 Local elections 13
7.4 Sugar quality 17
7.5 Physical training 22

7.1 Passing proportions

|||| Exercise 7.1 Passing proportions

To compare the level of 2 different courses at a university, the following grade distributions were recorded, showing the number of students who received each grade:

	Course 1	Course 2	*Row total*
Grade 12	20	14	34
Grade 10	14	14	28
Grade 7	16	27	43
Grade 4	20	22	42
Grade 2	12	27	39
Grade 0	16	17	33
Grade -3	10	22	32
Column total	108	143	251

The passing proportions for the two courses, p_1 and p_2, should be compared. As the grades -3 and 0 mean not passed, we get the following table of the number of students:

	Course 1	Course 2	*Row total*
Passed	82	104	186
Not passed	26	39	65
Column total	108	143	251

a) Compute a 95% confidence interval for the difference between the two passing proportions.

   Use Python to compute quantiles and probabilities from relevant probability distribution(s).

Solution

We use the formula for a large-sample confidence interval for the difference of two proportions (Method 7.15 in the textbook, p. 285; see also "Confidence interval for the difference" in the formula compendium, Section 7, p. 31):

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

with $x_1 = 82$, $n_1 = 108$, $x_2 = 104$, $n_2 = 143$ and $\alpha = 0.05$.

First we estimate the two proportions p_1 and p_2:

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{82}{108} = 0.759, \qquad \hat{p}_2 = \frac{x_2}{n_2} = \frac{104}{143} = 0.727.$$

The difference is therefore

$$\hat{p}_1 - \hat{p}_2 = \frac{82}{108} - \frac{104}{143} = 0.0320,$$

and the standard error of this difference is

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{(82/108)(26/108)}{108} + \frac{(104/143)(39/143)}{143}} = 0.0555.$$

We also need the quantile $z_{0.975}$, which we find using Python:

```
print(stats.norm.ppf(0.975))
```

```
1.959963984540054
```

Therefore the 95% confidence interval for $p_1 - p_2$ is

$$0.0320 \pm 1.96 \cdot 0.05549 = [-0.0768;\ 0.1408].$$

The interval contains 0, so the data do not show a statistically significant difference between the passing proportions of the two courses.

b) Use the z-test for comparing two proportions to assess whether the passing proportions differ between the two courses. The hypotheses are

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 \neq p_2.$$

What is the p-value, and what do you conclude at the significance level $\alpha = 0.01$?

Use Python to compute the p-value.

Solution

We use the z-test for comparing two proportions (textbook, Method 7.18, p. 287). The formula compendium gives the z-test statistic under "Two-proportions" (Section 7, p. 32).

Under H_0, the two courses have the same passing proportion, so we estimate this common proportion by pooling the counts from both courses:

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{82 + 104}{108 + 143} = \frac{186}{251} = 0.741.$$

Using the observed difference $\hat{p}_1 - \hat{p}_2 = 0.032$, the test statistic is then:

$$z_{\text{obs}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.032}{\sqrt{0.741 \cdot 0.259 \cdot \left(\frac{1}{108} + \frac{1}{143}\right)}} = \frac{0.032}{0.05585} = 0.573.$$

The test is two-sided, so the p-value is $2P(Z > |z_{\text{obs}}|)$, where $Z \sim N(0,1)$. The test statistic gives:

```
print(f"{2 * (1 - stats.norm.cdf(0.573)):.4f}")
```

```
0.5666
```

We can also carry out the test directly in Python using `proportions_ztest`, as shown in the formula compendium under "Two-proportions" (p. 32).

```
z_obs, p_val = smprop.proportions_ztest([82, 104], [108,
    143])
print(f"{p_val:.4f}")
```

```
0.5668
```

Since the p-value is approximately 0.567, which exceeds $\alpha = 0.01$, we cannot reject H_0. The data do not provide sufficient evidence that the passing proportions differ between the two courses. This is consistent with part a), where the 95% confidence interval for $p_1 - p_2$ contains 0.

c) Test the same hypothesis as in b), again with $\alpha = 0.01$, but now using the χ^2-test (comparing the same two proportions). What are the p-value and the conclusion? Compare with the result in b).

Use Python to compute quantiles and probabilities from relevant probability distribution(s).

|||| Solution

We use the χ^2-test for comparing proportions (textbook, Method 7.20, p. 290). The formulas are also given under "Contingency-tables" in the formula compendium (Section 7, p. 33).

We first find the expected number in each cell under H_0,

$$e_{ij} = \frac{\text{“row total”} \cdot \text{“column total”}}{\text{“grand total”}},$$

for example $e_{11} = \frac{186 \cdot 108}{251} = 80.03$. Doing this for all four cells gives:

	Course 1	Course 2
Passed	$o_{11} = 82$ $e_{11} = 80.03$	$o_{12} = 104$ $e_{12} = 105.97$
Not passed	$o_{21} = 26$ $e_{21} = 27.97$	$o_{22} = 39$ $e_{22} = 37.03$

All expected counts exceed 5, so the rule of thumb for using the χ^2-approximation is satisfied.

The test statistic is the sum of the four contributions:

$$\begin{aligned}\chi^2_{\text{obs}} &= \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(o_{ij}-e_{ij})^2}{e_{ij}} \\ &= \frac{(82-80.03)^2}{80.03} + \frac{(104-105.97)^2}{105.97} + \frac{(26-27.97)^2}{27.97} + \frac{(39-37.03)^2}{37.03} \\ &= 0.048 + 0.037 + 0.138 + 0.105 = 0.328.\end{aligned}$$

The degrees of freedom are $(2-1)(2-1) = 1$, so the p-value is $P(\chi^2 > 0.328)$ in the $\chi^2(1)$-distribution:

```
print(f"{1 - stats.chi2.cdf(0.328, 1):.4f}")
```

```
0.5668
```

The test can also be carried out directly in Python using `proportions_chisquare`, as shown in the formula compendium under "Contingency-tables" (p. 33).

```
chi2_obs, p_val, tables = smprop.proportions_chisquare([82,
    104], [108, 143])
print(f"{p_val:.4f}")
```

```
0.5668
```

The p-value 0.567 is much larger than $\alpha = 0.01$, so again we cannot reject H_0. The data does not provide sufficient evidence that the passing proportions differ between the two courses.

Comparison with b). The p-value is the same as in b), apart from rounding. For a 2×2 table, the two-sided z-test and the χ^2-test used here are equivalent. Their test statistics satisfy

$$\chi^2_{\text{obs}} = (z_{\text{obs}})^2 = 0.573^2 = 0.328,$$

as given in the textbook, Equation (7-51), p. 293.

d) For a course offered repeatedly, the passing proportion is assumed to be 0.80 on average, and 250 students take the exam each time. What are the expected value, μ, and the standard deviation, σ, of the number of students who do not pass the exam?

|||| Solution

If X is the number of students who do not pass the exam on one occasion, then, assuming independent outcomes, this random variable follows the binomial distribution with $n = 250$ and $p = 0.20$, so we use the formulas from Theorem 2.21 (p. 58) for the mean and variance of the binomial distribution. These are also given under "Binomial distribution" in the formula compendium (Section 2, p. 4).

$$\mu = np = 250 \cdot 0.20 = 50, \qquad \sigma^2 = np(1-p) = 250 \cdot 0.20 \cdot 0.80 = 40.$$

The standard deviation is therefore

$$\sigma = \sqrt{\sigma^2} = \sqrt{40} = 6.32.$$

Thus the expected number of students who do not pass is 50, with a standard deviation of 6.32 students.

Exam-style question

e) What is the critical value for the χ^2-test in part c) at the significance level $\alpha = 0.01$?

1) 2.576
2) 3.841
3) 6.635
4) 7.879
5) 9.210

|||| Solution

The 2×2 table gives $(2-1)(2-1) = 1$ degree of freedom. The χ^2-test rejects H_0 for large values of the test statistic, so the critical value is the 0.99 quantile of the $\chi^2(1)$ distribution (formula compendium, Section 7, p. 33).

```
print(f"{stats.chi2.ppf(0.99, 1):.3f}")
```

```
6.635
```

The correct answer is therefore **3)**.

7.2 Outdoor lighting

Exercise 7.2 Outdoor lighting

A company sells an outdoor lamp in three variants: copper, painted and stainless steel. Some of the lamps are sold in Denmark, and the rest are exported. The table shows the percentage of all 250 lamps sold in each combination of variant and sales destination.

	Sales Destination	
	Denmark	Export
Copper variant	7.2%	6.4%
Painted variant	28.0%	34.8%
Stainless steel variant	8.8%	14.8%

a) Consider all three variants together. Is there a significant difference between the proportion of the 250 lamps that is exported and the proportion that is sold in Denmark (at $\alpha = 0.05$)?

Use Python to compute quantiles and probabilities from relevant probability distribution(s).

Solution

We use a one-sample proportion test. Adding the percentages in the Denmark column gives $7.2\% + 28.0\% + 8.8\% = 44\%$, so $0.44 \cdot 250 = 110$ lamps were sold in Denmark and the remaining $250 - 110 = 140$ were exported.

Given that p is the proportion of lamps exported, under the null hypothesis we would expect the same number to be sold in Denmark as to be exported: 125 of the 250 lamps in each group. This corresponds to $p = 0.5$, so we test:

$$H_0 : p = 0.5, \qquad H_1 : p \neq 0.5,$$

Using the one-sample proportion test (textbook, Method 7.11, p. 282; formula compendium, "One-proportion", p. 30), the test statistic is

$$z_{\text{obs}} = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{140 - 250 \cdot 0.5}{\sqrt{250 \cdot 0.5 \cdot 0.5}} = \frac{15}{\sqrt{250/4}} = \frac{15}{7.906} = 1.90.$$

For this two-sided test at $\alpha = 0.05$, the critical values are $\pm z_{0.975}$.

```
print(stats.norm.ppf(0.975))
```

```
1.959963984540054
```

Since $-1.96 < 1.90 < 1.96$, we do not reject H_0. The data do not show a statistically significant difference between the proportion exported and the proportion sold in Denmark.

b) Now take the variants into account. Is there a significant difference in how the variants are distributed in Denmark and for export (at $\alpha = 0.05$)? Carry out the relevant hypothesis test and report the test statistic, degrees of freedom, p-value, critical value and conclusion.

Hint: start by converting the table into counts instead of percentages.

Use Python to carry out the hypothesis test and to compute quantiles and probabilities from relevant probability distribution(s).

|||| Solution

We use a χ^2-test of homogeneity in a 3×2 frequency table (textbook, Method 7.22, p. 294). The formulas and Python command are also given under "Contingency-tables" in the formula compendium (Section 7, p. 33).

Under H_0, each lamp variant accounts for the same proportion of sales in Denmark as it does among exports. In symbols,

$$H_0 : p_{\text{variant, Denmark}} = p_{\text{variant, export}}$$

for each of the three variants. Under H_1, the proportions differ for at least one variant.

The observed counts. The χ^2-test uses counts, so we first multiply each percentage by the total number of lamps, 250. This gives

	Denmark	Export	*Row total*
Copper variant	18	16	34
Painted variant	70	87	157
Stainless steel variant	22	37	59
Column total	110	140	250

The χ^2-test in Python. We store the counts in an array and carry out the test using `stats.chi2_contingency`.

```
obs = np.array([[18, 16],
                [70, 87],
                [22, 37]])
chi2_obs, p_val, dof, exp = stats.chi2_contingency(obs)
```

The test statistic compares the observed counts o_{ij} with the expected counts e_{ij} under H_0,

$$\chi^2_{\text{obs}} = \sum_{i=1}^{3}\sum_{j=1}^{2}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}.$$

Its value is

```
print(chi2_obs)
```

```
2.2037040503932817
```

The degrees of freedom are $(r-1)(c-1) = (3-1)(2-1) = 2$.

```
print(dof)
```

```
2
```

The p-value is $P(\chi^2 > 2.204)$ in the $\chi^2(2)$-distribution.

```
print(p_val)
```

```
0.33225516858504067
```

The function also returns the expected counts under H_0, calculated as

$$e_{ij} = \frac{\text{row total } i \cdot \text{column total } j}{\text{grand total}}.$$

```
print(exp)
```

```
[[14.96 19.04]
 [69.08 87.92]
 [25.96 33.04]]
```

All expected counts are at least 5, so the rule of thumb for using the χ^2-approximation is satisfied.

The critical value. At $\alpha = 0.05$, the critical value is $\chi^2_{0.95}(2)$.

```
print(stats.chi2.ppf(0.95, 2))
```

```
5.991464547107979
```

Conclusion. The observed test statistic $\chi^2_{\text{obs}} = 2.204$ is smaller than the critical value 5.991, so we do not reject H_0 at $\alpha = 0.05$. We reach the same conclusion from the p-value, since $0.332 > 0.05$.

The data do not show a statistically significant difference in the distribution of lamp variants between sales in Denmark and export sales.

Exam-style question

c) Suppose the company had recorded sales of 500 lamps, with exactly the same percentages as in the table. How would the χ^2-test statistic and the p-value change compared with part b)?

1) The test statistic stays the same and the p-value stays the same.
2) The test statistic halves and the p-value increases.
3) The test statistic quadruples and the p-value decreases.
4) The test statistic doubles and the p-value decreases.
5) The test statistic doubles and the p-value stays the same.

|||| Solution

With twice as many lamps and the same percentages, each observed count doubles. Each expected count also doubles, since the row totals, column totals and grand total all double. Using the χ^2-test statistic from part b) (formula compendium, Section 7, p. 33), the contribution from each cell becomes

$$\frac{(2o_{ij} - 2e_{ij})^2}{2e_{ij}} = 2\frac{(o_{ij} - e_{ij})^2}{e_{ij}}.$$

The test statistic therefore doubles from 2.204 to 4.407 (rounded). The table still has two degrees of freedom, so the larger test statistic gives a smaller upper-tail probability. The p-value decreases from 0.332 to 0.110, so the correct answer is **4)**.

7.3 Local elections

Exercise 7.3 Local elections

At the local elections in Denmark in November 2013, the Social Democrats (A) received $p = 29.5\%$ of the votes at the country level. An early exit poll (a survey of voters leaving polling stations) estimated that they would get 22.7% of the votes. Suppose the exit poll was based on 740 people, out of which 168 people reported having voted for A.

a) At the time of the exit poll, the nationwide share of the votes, p, was of course not known. Did the poll indicate that this share would differ significantly from 29.5%? Test the hypotheses:

$$H_0 : p = 0.295$$
$$H_1 : p \neq 0.295.$$

What is the test statistic, and what do you conclude at the significance level $\alpha = 0.001$?

Use Python to compute quantiles and probabilities from relevant probability distribution(s).

Solution

We use the one-sample proportion test (Method 7.11 in the textbook, p. 282). The test statistic and critical values are also given under "One-proportion" in the formula compendium (Section 7, p. 30). With $x = 168$, $n = 740$ and $p_0 = 0.295$, the test statistic is

$$z_{\text{obs}} = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{168 - 740 \cdot 0.295}{\sqrt{740 \cdot 0.295 \cdot (1 - 0.295)}} = \frac{168 - 218.3}{\sqrt{153.9}} = \frac{-50.3}{12.41} = -4.05.$$

For this two-sided test at $\alpha = 0.001$, the critical values are $\pm z_{0.9995}$.

```
print(stats.norm.ppf(0.9995))
```

```
3.2905267314919255
```

Since $z_{\text{obs}} = -4.05$ is below the lower critical value of -3.291, we reject H_0. Based on the exit poll, we would have concluded that A's nationwide share of the votes differed significantly from 29.5%, even though A actually received 29.5%.

The test assumes that the exit poll respondents form a random sample of all voters. This is hard to achieve in practice. For example, some people refuse to answer, and an early poll misses those who vote later. The large difference from the election result suggests that this exit poll was not representative.

We can reach the same conclusion from the two-sided p-value, $2 \cdot P(Z < -4.05)$, where $Z \sim N(0,1)$.

```
print(2 * stats.norm.cdf(-4.05))
```

```
5.121763294808298e-05
```

This is much smaller than $\alpha = 0.001$.

b) Calculate a 95%-confidence interval for p based on the exit poll.

Use Python to compute quantiles and probabilities from relevant probability distribution(s).

|||| Solution

The exit poll gives the following estimate of A's nationwide share of the votes:

$$\hat{p} = \frac{x}{n} = \frac{168}{740} = 0.227,$$

We use the confidence interval for one proportion from Method 7.3 in the textbook (p. 277), also given under "One-proportion" in the formula compendium (Section 7, p. 29):

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

The estimated standard error is:

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.227 \cdot 0.773}{740}} = 0.0154,$$

For a 95% confidence interval, we need the quantile $z_{0.975}$.

```
print(stats.norm.ppf(0.975))
```

```
1.959963984540054
```

The 95% confidence interval is therefore

$$0.227 \pm 1.96 \cdot 0.0154 = 0.227 \pm 0.030 = [0.197;\ 0.257].$$

Notice that the interval does not contain the election result 0.295, in agreement with the rejection in a). This means that, at the time of the exit poll, we would have said that the plausible values for A's share of the votes were between 19.7% and 25.7%. Again, this assumes that the exit poll is a random sample of all voters.

c) Based on a scenario in which the proportion voting for a particular party is around 30%, how large should an exit poll be to achieve a 99% confidence interval having a width of 0.01 on average for this proportion?

Use Python to compute quantiles and probabilities from relevant probability distribution(s).

|||| Solution

We use $p = 0.30$ as a guess for the unknown proportion, based on the assumption in the question that around 30% of voters will vote for the party. The sample size formula from Method 7.13 in the textbook (p. 284) is also given under "One-proportion" in the formula compendium (Section 7, p. 30). It gives

$$n = 0.30 \cdot 0.70 \left(\frac{z_{0.995}}{ME}\right)^2,$$

where ME is the margin of error. Since the confidence interval extends one margin of error on each side of the estimate, its total width is twice the margin of error. We therefore take

$$ME = \frac{0.01}{2} = 0.005.$$

For a 99% confidence interval, $\alpha = 0.01$, so the required quantile is $z_{1-\alpha/2} = z_{0.995}$.

```
print(stats.norm.ppf(0.995))
```

```
2.5758293035489004
```

Using the rounded value $z_{0.995} = 2.576$ gives

$$n = 0.30 \cdot 0.70 \left(\frac{2.576}{0.005} \right)^2 = 0.21 \cdot 515.2^2 = 0.21 \cdot 265431 = 55740.5.$$

Rounding up to the nearest integer gives a required sample size of 55741 voters.

Using the unrounded quantile gives $n = 55733.1$, which rounds up to 55734 voters. The difference comes from rounding $z_{0.995}$. Again, this assumes that the exit poll respondents form a random sample of all voters.

Exam-style question

d) Suppose no prior estimate of the party's share of the votes is available. What is the minimum sample size needed to plan an exit poll with a 95% confidence interval of width at most 0.02, for any value of p? Use $z_{0.975} = 1.96$.

1) 2 401
2) 9 604
3) 8 068
4) 16 588
5) 38 416

|||| Solution

The required sample size is largest when $p(1-p)$ is largest, which occurs at $p = 0.5$. We therefore use $p = 0.5$ when no prior estimate is available. The margin of error is half the total width, so $ME = 0.02/2 = 0.01$.

Using the sample size formula from part c) (formula compendium, Section 7, p. 30), we obtain

$$n = 0.5 \cdot 0.5 \left(\frac{1.96}{0.01} \right)^2 = 9\,604.$$

The exit poll therefore requires at least 9 604 respondents. The correct answer is **2)**.

7.4 Sugar quality

|||| Exercise 7.4 Sugar quality

A wholesaler needs to find a supplier that delivers sugar in 1 kg bags. The wholesaler receives 50 bags from each of two potential suppliers. A bag is considered 'defective' if the weight of the filled bag is less than 990 grams. All the received bags were weighed, and 6 defective bags from supplier A and 12 from supplier B were found.

a) Given that p_A and p_B denote the proportions of defective bags from suppliers A and B, respectively, test whether these proportions differ. The hypotheses being considered are:

$$H_0 : p_A = p_B$$
$$H_1 : p_A \neq p_B.$$

What are the p-value and the conclusion at a significance level of 5%?

Use Python to compute quantiles and probabilities from relevant probability distribution(s).

|||| Solution

We use the z-test for comparing two proportions (textbook, Method 7.18, p. 287). The test statistic is also given under "Two-proportions" in the formula compendium (Section 7, p. 32). The test uses a normal approximation and assumes independent random samples from the two suppliers.

First we estimate the two proportions of defective bags:

$$\hat{p}_A = \frac{x_A}{n_A} = \frac{6}{50} = 0.12, \qquad \hat{p}_B = \frac{x_B}{n_B} = \frac{12}{50} = 0.24.$$

Under H_0, the two suppliers have the same proportion of defective bags, so we estimate this common proportion by pooling the counts from both samples:

$$\hat{p} = \frac{x_A + x_B}{n_A + n_B} = \frac{6 + 12}{50 + 50} = 0.18.$$

The test statistic is then

$$z_{\text{obs}} = \frac{\hat{p}_B - \hat{p}_A}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} = \frac{0.24 - 0.12}{\sqrt{0.18 \cdot 0.82 \cdot (2/50)}} \approx \frac{0.12}{0.07684} \approx 1.5617.$$

The test is two-sided, so the p-value is $2P(Z > |z_{\text{obs}}|)$, where $Z \sim N(0,1)$. We keep the unrounded test statistic when calculating this probability in Python.

```
z_obs = (0.24 - 0.12) / np.sqrt(0.18 * 0.82 * (2 / 50))
print(f"{2 * (1 - stats.norm.cdf(abs(z_obs))):.4f}")
```

```
0.1183
```

The test can also be carried out directly in Python:

```
z_obs, p_val = smprop.proportions_ztest([6, 12], [50, 50])
print(f"{p_val:.4f}")
```

```
0.1183
```

Here, `z_obs` is approximately -1.5617, because the function uses $\hat{p}_A - \hat{p}_B$. This does not change the two-sided p-value.

The p-value 0.118 is larger than $\alpha = 0.05$, so we do not reject H_0. The data do not provide sufficient evidence that the proportions of defective bags differ between the two suppliers.

b) A supplier has delivered 200 bags, of which 36 were defective. Find a 99% confidence interval for p, the proportion of defective bags for this supplier.

Use Python to compute quantiles and probabilities from relevant probability distribution(s).

|||| Solution

We use the confidence interval for one proportion (textbook, Method 7.3, p. 277), also given under "One-proportion" in the formula compendium (Section 7, p. 29):

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

with $\alpha = 0.01$. The estimate is

$$\hat{p} = \frac{x}{n} = \frac{36}{200} = 0.18,$$

and its estimated standard error is

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.18 \cdot 0.82}{200}} \approx 0.0272.$$

For a 99% confidence interval, the required quantile is $z_{1-\alpha/2} = z_{0.995}$.

```
print(stats.norm.ppf(0.995))
```

```
2.5758293035489004
```

Using these rounded values, the 99% confidence interval is approximately

$$0.18 \pm 2.576 \cdot 0.0272 \approx 0.18 \pm 0.070 = [0.110;\ 0.250].$$

The interval can also be found directly in Python:

```
lower, upper = smprop.proportion_confint(36, 200, alpha
    =0.01)
print(f"({lower:.3f}, {upper:.3f})")
```

```
(0.110, 0.250)
```

Therefore, the 99% confidence interval for the proportion of defective bags from this supplier is approximately $[0.110;\ 0.250]$, or 11% to 25%, assuming that the bags form a random sample from the supplier

c) A new study was planned on the assumption that about 20% of the bags from a new supplier are defective.The aim was to obtaining an average width, B, of a 95% confidence interval. The Analysis Department calculated that 1537 bags should be examined, but forgot to record the value of B. What value of B was used?

Use Python to compute quantiles and probabilities from relevant probability distribution(s).

|||| Solution

We use $p = 0.20$, as assumed in the question. The sample size formula from Method 7.13 in the textbook (p. 284) is also given under "One-proportion" in the formula compendium (Section 7, p. 30):

$$n = p(1-p)\left(\frac{z_{0.975}}{ME}\right)^2 = 0.20 \cdot 0.80\left(\frac{z_{0.975}}{ME}\right)^2,$$

where ME is the margin of error. Here $n = 1537$ is known and ME is unknown, so we solve for ME:

$$ME = \sqrt{\frac{0.20 \cdot 0.80 \cdot z_{0.975}^2}{n}}.$$

For a 95% confidence interval, the required quantile is $z_{0.975}$.

```
print(stats.norm.ppf(0.975))
```

```
1.959963984540054
```

Substituting $n = 1537$ and using $z_{0.975} \approx 1.96$ gives

$$ME \approx \sqrt{\frac{0.16 \cdot 1.96^2}{1537}} \approx 0.020.$$

Since the width of the confidence interval is twice the margin of error, the value used by the Analysis Department was approximately

$$B = 2ME \approx 0.040,$$

corresponding to a width of 4 percentage points.

Exam-style question

d) Based on the test in a), the wholesaler concludes: "The two suppliers deliver the same proportion of defective bags, so it does not matter which one we choose."
Is the wholesaler's conclusion justified? Choose the option in which both the judgment and the reason are correct.

1) It is not justified, because failing to reject H_0 does not establish that the proportions are equal.
2) It is justified, because a p-value above 0.05 means that the proportions are equal with 95% confidence.
3) It is not justified, because the observed proportions 0.12 and 0.24 prove that the true proportions differ.
4) It is justified, because the p-value of 0.118 is the probability that the two proportions differ.
5) It is justified, because a confidence interval for $p_A - p_B$ that contains zero proves that the difference is zero.

Solution

The correct answer is **1)**. Since the p-value of 0.118 exceeds 0.05, we do not reject H_0. The data provide insufficient evidence that the proportions differ. A real difference could go undetected with 50 bags from each supplier, so the test does not establish that the proportions are equal (textbook, Remark 3.26, p. 137).

7.5 Physical training

Exercise 7.5 Physical training

A company wants to investigate whether the employees' physical training condition is associated with their job success. The following table shows the results for the 200 employees that were tested

	Physical training condition		
	Below average	Average	Above average
Bad job success	11	27	15
Average job success	14	40	30
Good job success	5	23	35

We use a χ^2-test to test the null hypothesis H_0 that job success and physical training condition are independent.

a) Assuming that H_0 is true, what is the expected number of employees with an above average training condition and good job success?

Solution

We use the formula for expected cell counts under H_0 (textbook, Eq. (7-53), p. 294), also given under "Contingency-tables" in the formula compendium (Section 7, p. 33). For cell (i, j), this gives

$$e_{ij} = \frac{(\text{row total } i)(\text{column total } j)}{\text{grand total}}.$$

To apply this formula, we first calculate the row and column totals. We can do so in Python:

```
obs = np.array([[11, 27, 15],
                [14, 40, 30],
                [ 5, 23, 35]])
print(obs.sum(axis=1))  # row totals
print(obs.sum(axis=0))  # column totals
```

```
[53  84  63]
[30  90  80]
```

Under H_0, we expect the overall proportion with good job success, $63/200$, to apply to the 80 employees with above-average training condition. The expected count in cell $(3,3)$ is therefore

$$e_{33} = 80 \cdot \frac{63}{200} = 25.2.$$

b) For the calculation of the relevant χ^2-test statistic, identify the following two numbers:
   - A: the number of contributions to the test statistic
   - B: the contribution to the statistic from table cell $(1,1)$

|||| Solution

We use the χ^2-test statistic for an $r \times c$ frequency table (textbook, Method 7.22, p. 294), also given under "Contingency-tables" in the formula compendium (Section 7, p. 33):

$$\chi^2_{\text{obs}} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(o_{ij} - e_{ij})^2}{e_{ij}},$$

where o_{ij} and e_{ij} are the observed and expected counts in cell (i,j).

Each cell contributes one term to the sum. Since the table has 3 rows and 3 columns, the number of contributions is

$$A = 3 \cdot 3 = 9.$$

For cell $(1,1)$, the observed count is $o_{11} = 11$. Using the row total 53 and column total 30, the expected count under H_0 is

$$e_{11} = \frac{53 \cdot 30}{200} = 7.95.$$

The contribution from this cell is therefore

$$B = \frac{(o_{11} - e_{11})^2}{e_{11}} = \frac{(11 - 7.95)^2}{7.95} \approx 1.170.$$

c) The total χ^2-test statistic is 10.985. What are the p-value and the conclusion at the 5% significance level?

Use Python to compute quantiles and probabilities from relevant probability distribution(s).

|||| Solution

We use the reference distribution for the χ^2-test statistic (textbook, Method 7.22, p. 294), also given under "Contingency-tables" in the formula compendium (Section 7, p. 33). The degrees of freedom are

$$(r-1)(c-1) = (3-1)(3-1) = 4.$$

Under H_0, the test statistic approximately follows a χ^2 distribution with 4 degrees of freedom. Larger values indicate greater discrepancies between the observed and expected counts, so the p-value is the upper-tail probability

$$P(\chi^2_4 \geq 10.985).$$

We calculate this probability in Python.

```
p_val = 1 - stats.chi2.cdf(10.985, df=4)
print(f"{p_val:.4f}")
```

```
0.0267
```

Since $0.0267 < 0.05$, we reject H_0 at the 5% significance level. The data provide evidence of an association between job success and physical training condition.

The full test can also be carried out directly from the observed counts.

```
chi2_obs, p_val, dof, exp = stats.chi2_contingency(obs)
print(f"{chi2_obs:.4f} {p_val:.4f} {dof}")
```

```
10.9849 0.0267 4
```

Exam-style question

d) Under H_0, the expected count in cell $(3,3)$ is $e_{33} = 25.2$. The observed count is $o_{33} = 35$. What is the contribution of this cell to the χ^2-test statistic?

1) 2.744
2) 3.811
3) 0.389
4) 96.04
5) 9.333

|||| Solution

We use the formula for the χ^2-test statistic (textbook, Method 7.22, Eq. (7-55), p. 294), also given under "Contingency-tables" in the formula compendium (Section 7, p. 33).

Each cell contributes the squared difference between its observed and expected counts, divided by its expected count. For cell $(3,3)$, this gives

$$\frac{(o_{33} - e_{33})^2}{e_{33}} = \frac{(35 - 25.2)^2}{25.2} \approx 3.811.$$

The correct answer is **2)**.