Comparing Proportions

Last modified: 2026-11-12 22:02:32 (PST)

This page reviews tests for comparing groups on a categorical outcome: the chi-square test and Fisher's exact test for contingency tables. It uses the chi-square reference distribution defined on the Statistical Inference page. This page is adapted from Vittinghoff et al. [1], Chapter 3.

The HERS data

The examples on this page use the HERS data, which the Comparing Means page describes. The `rmb` R package includes the dataset; `haven::as_factor()` converts its Stata value labels to factors:

```
hers <- rmb::hers |> haven::as_factor()
hers |> head()
#> # A tibble: 6 x 37
#>   HT      age raceth nonwhite smoking drinkany exercise physact globrat poorfair
#>   <fct> <dbl> <fct>  <fct>    <fct>   <fct>    <fct>    <fct>   <fct>   <fct>
#> 1 plac~    70 Afric~ yes      no      no       no       much m~ good    no
#> 2 plac~    62 Afric~ yes      no      no       no       much l~ good    no
#> 3 horm~    69 White  no       no      no       no       about ~ good    no
#> 4 plac~    64 White  no       yes     yes      no       much l~ good    no
#> 5 plac~    65 White  no       no      no       no       somewh~ good    no
#> 6 horm~    68 Afric~ yes      no      yes      no       about ~ good    no
#> # i 27 more variables: medcond <dbl>, htnmeds <fct>, statins <fct>,
#> #   diabetes <fct>, dmpills <fct>, insulin <fct>, weight <dbl>, BMI <dbl>,
#> #   waist <dbl>, WHR <dbl>, glucose <dbl>, weight1 <dbl>, BMI1 <dbl>,
#> #   waist1 <dbl>, WHR1 <dbl>, glucose1 <dbl>, tchol <dbl>, LDL <dbl>,
#> #   HDL <dbl>, TG <dbl>, tchol1 <dbl>, LDL1 <dbl>, HDL1 <dbl>, TG1 <dbl>,
#> #   SBP <dbl>, DBP <dbl>, age10 <dbl>
```

Comparing two groups: categorical outcomes

Contingency tables

Example 0.1 (Exercise by treatment group in HERS)**.** Table 1 cross-tabulates regular exercise at baseline by treatment group, as a contingency table with row percentages.

Table 1: Exercise by treatment group in HERS

```
hers |>
  gtsummary::tbl_cross(
    row = exercise,
    col = HT,
    label = list(
      exercise ~ "Exercises regularly",
      HT ~ "Treatment group"
    ),
    percent = "row"
  )
```

	Treatment group		
	placebo	hormone therapy	Total
Exercises regularly			
no	853 (50%)	842 (50%)	1,695 (100%)
yes	530 (50%)	538 (50%)	1,068 (100%)
Total	1,383 (50%)	1,380 (50%)	2,763 (100%)

Read

- Hogg et al. [2, Sec. 7.3], pp. 323–328 covers confidence intervals for proportions.
- Hogg et al. [2, Sec. 8.4], pp. 385–391 covers tests about proportions.

The chi-square test

Definition 0.1 (Expected count under independence)**.** Let a contingency table have r rows and c columns, with observed count O_{ij} in row i and column j, row totals R_i, column totals C_j, and grand total n. The **expected count** in cell (i, j) under independence is

$$E_{ij} \stackrel{\text{def}}{=} \frac{R_i\, C_j}{n}.$$

Source: Wikipedia contributors [3].

Example 0.2 (Expected count in a 2 x 2 table)**.** In a table with $n = 100$, first-row total $R_1 = 30$, and first-column total $C_1 = 40$, the expected count in cell $(1, 1)$ is

$$\begin{aligned} E_{11} &= 30 \cdot 40/100 \\ &= 12. \end{aligned}$$

Definition 0.2 (Pearson's chi-square test of independence)**.** With observed counts O_{ij} and expected counts E_{ij} in an $r \times c$ contingency table, **Pearson's chi-square test** of the null hypothesis that the row and column variables are independent uses the statistic

$$X^2 \stackrel{\text{def}}{=} \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}.$$

Its p-value is $\Pr(W \geq X^2)$, where W has the $\chi^2_{(r-1)(c-1)}$ distribution (chi-square distribution).
Source: Wikipedia contributors [4].

Theorem 0.1 (Large-sample null distribution of the chi-square statistic)**.** *Let n observations be sampled independently and classified by two categorical variables, and let the two variables be independent. Then as $n \to \infty$, the distribution of X^2 (Definition 0.2) converges to the $\chi^2_{(r-1)(c-1)}$ distribution [2, Sec. 9.2], p. 440.*

ℹ Remark (Small expected counts)

Remark 0.1 (Small expected counts). The chi-square approximation is poor when some expected counts are small. A common rule of thumb asks for every E_{ij} to be at least 5.

Example 0.3 (Chi-square test of exercise by treatment group in HERS)**.** The expected counts and statistic of Definition 0.2 for Table 1:

```r
observed <- table(hers$exercise, hers$HT)
expected <- outer(rowSums(observed), colSums(observed)) / sum(observed)
x_squared <- sum((observed - expected)^2 / expected)
expected
#>     placebo hormone therapy
#> no   848.42          846.58
#> yes  534.58          533.42
c(
  x_squared = x_squared,
  p_value = pchisq(x_squared, df = 1, lower.tail = FALSE)
)
#> x_squared   p_value
#>  0.128054  0.720458
```

Every expected count is large, so the χ^2_1 approximation of Theorem 0.1 is reasonable. `chisq.test()` with `correct = FALSE` reports the same values:

```r
chisq.test(hers$exercise, hers$HT, correct = FALSE)
#>
#>  Pearson's Chi-squared test
#>
#> data:  hers$exercise and hers$HT
#> X-squared = 0.1281, df = 1, p-value = 0.72
```

The p-value is large: the data give no evidence that exercise depends on treatment group, as randomization would lead us to expect.

Remark (Yates' continuity correction)

Remark 0.2 (Yates' continuity correction). For a 2×2 table, `chisq.test()` applies Yates' continuity correction by default, which subtracts 0.5 from each $|O_{ij} - E_{ij}|$ before squaring, and so gives a smaller statistic than Definition 0.2. `correct = FALSE` turns the correction off.

Read

- Hogg et al. [2, Sec. 9.1], pp. 425–434 covers chi-square goodness-of-fit tests.
- Hogg et al. [2, Sec. 9.2], pp. 435–445 covers contingency tables.
- Rice [5, Sec. 13.3], pp. 516–519 covers the chi-square test of homogeneity.
- Rice [5, Sec. 13.4], pp. 520–522 covers the chi-square test of independence.
- Rice [5, Sec. 9.5], pp. 341–346 covers likelihood ratio tests for the multinomial

distribution.

Fisher's exact test

Definition 0.3 (Fisher's exact test)**.** Take a 2×2 contingency table with cells a, b, c, and d as in the contingency table definition, and hold its row and column totals fixed. Under independence of the row and column variables, the probability that the top-left cell equals x is the hypergeometric probability

$$p(x) \stackrel{\text{def}}{=} \frac{\binom{a+b}{x}\binom{c+d}{a+c-x}}{\binom{n}{a+c}}.$$

Fisher's exact test of independence has p-value [6]

$$\sum_{x\,:\,p(x)\le p(a)} p(x),$$

the total probability of the tables with the same totals that are no more probable than the observed table.

Remark (Why Fisher's exact test is exact)

Remark 0.3 (Why Fisher's exact test is exact). The p-value is exact: it comes from the null distribution itself, not from a large-sample approximation, so the test is valid even when expected counts are small, and it is often used for 2×2 tables in which some expected count is below 5 (Theorem 0.1). Other two-sided versions exist; this one is the version that R's `fisher.test()` computes.

Example 0.4 (Fisher's exact test of exercise by treatment group in HERS)**.** The p-value of Definition 0.3 for Table 1, computed from the hypergeometric probabilities with `dhyper()`:

```
a <- observed[1, 1]
row1 <- sum(observed[1, ])
row2 <- sum(observed[2, ])
col1 <- sum(observed[, 1])
x <- max(0, col1 - row2):min(row1, col1)
p_x <- dhyper(x, m = row1, n = row2, k = col1)
p_a <- dhyper(a, m = row1, n = row2, k = col1)
sum(p_x[p_x <= p_a * (1 + 1e-7)])
#> [1] 0.72528
```

The tolerance `1e-7` keeps tables whose probability equals $p(a)$ up to rounding error, as `fisher.test()` does. `fisher.test()` reports the same p-value:

```
fisher.test(hers$exercise, hers$HT)
#>
#>  Fisher's Exact Test for Count Data
#>
#> data:  hers$exercise and hers$HT
#> p-value = 0.725
#> alternative hypothesis: true odds ratio is not equal to 1
#> 95 percent confidence interval:
#>  0.879664 1.202192
#> sample estimates:
#> odds ratio
#>    1.02836
```

With counts this large, the exact p-value is close to the chi-square p-value of Example 0.3.

Read

Rice [5, Sec. 13.2], pp. 514–515 covers Fisher's exact test.

Measures of association for 2×2 tables

Tests of independence say whether two binary variables are associated, but not how strongly. The risk difference, the risk ratio and the odds ratio measure the strength of the association.

Definition 0.4 (Risk of a binary outcome (also called the incidence proportion or cumulative incidence))**.** The **risk** of a binary outcome in a group is the probability that a member of the group has the outcome. When the outcome is an event, the risk is for a specified time period. The risk is also called the **incidence proportion** or **cumulative incidence**.
The risk lies in $[0, 1]$, and its maximum likelihood estimate is the proportion of the group's observations that have the outcome.
Source: Wikipedia contributors [7].

Example 0.5 (Risk of regular exercise in each HERS treatment group)**.** In Table 1, the outcome is regular exercise at baseline. The estimated risk in each treatment group is the proportion of that group's column who exercise regularly:

```r
risk_hat <- prop.table(table(hers$exercise, hers$HT), margin = 2)["yes", ]
risk_hat
#>         placebo hormone therapy
#>        0.383225        0.389855
```

The estimated risk of regular exercise is 0.383 in the placebo group and 0.39 in the hormone therapy group.

Remark (Cumulative incidence has a second meaning)

Remark 0.4 (Cumulative incidence has a second meaning). The name cumulative incidence for the risk fits a closed population, where no one enters or leaves after the start of the time period [8, Ch. 4]. The same name is also used for a different quantity: when several event types compete, the cumulative incidence function of one event type at time t is the probability that this type of event occurs by time t [8, Ch. 22]. This page uses the name only for the risk of a binary outcome.

Definition 0.5 (Reference group)**.** In a comparison of an outcome measure between two groups, the **reference group** is the group whose measure is the subtrahend of a difference or the denominator of a quotient. The other group is the group being compared with it.

Example 0.6 (The placebo group as the reference group in HERS)**.** To compare the risks of regular exercise in Example 0.5, take the placebo group as the reference group and the hormone therapy group as the group being compared with it. Then a difference is hormone therapy minus placebo, and a quotient is hormone therapy divided by placebo.

Definition 0.6 (Risk difference (also called the absolute risk reduction))**.** Let:

- π_1 be the risk of a binary outcome in group 1.
- π_2 be the risk of the same outcome in the reference group, group 2.

The **risk difference**, also called the **absolute risk reduction**, of group 1 compared with group 2 is

$$\mathrm{RD} \stackrel{\text{def}}{=} \pi_1 - \pi_2 \in [-1, 1].$$

Source: Wikipedia contributors [9].

Example 0.7 (Risk difference of regular exercise in HERS)**.** Take group 1 to be hormone therapy and group 2, the reference group, to be placebo. The estimated risk difference for the estimated risks of Example 0.5 is:

```
rd_hat <- risk_hat[["hormone therapy"]] - risk_hat[["placebo"]]
rd_hat
#> [1] 0.0066302
```

The estimated risk of regular exercise is 0.007 higher in the hormone therapy group than in the placebo group. Multiplying by 100 gives 0.7 percentage points.

Exercise 0.1. Let:

- π_1 and π_2 be the risks of a binary outcome in group 1 and group 2.
- $\text{RD} \stackrel{\text{def}}{=} \pi_1 - \pi_2$ be the risk difference of group 1 compared with group 2.

Express the risk difference of group 2 compared with group 1, $\pi_2 - \pi_1$, in terms of RD.

Solution

Solution 0.1.

$$\begin{aligned} \pi_2 - \pi_1 &= -1 \cdot (\pi_1 - \pi_2) && \text{(factor } -1 \text{ out of both terms)} \\ &= -\,\text{RD} && \text{(definition of RD)} \end{aligned}$$

Theorem 0.2 (Switching the reference group in a risk difference)**.** *Let:*

- π_1 *and* π_2 *be the risks of a binary outcome in group 1 and group 2.*
- $\text{RD} \stackrel{\text{def}}{=} \pi_1 - \pi_2$ *be the risk difference of group 1 compared with group 2, with group 2 as the reference group.*

Then the risk difference of group 2 compared with group 1, with group 1 as the reference group, is the negative of RD*:*

$$\pi_2 - \pi_1 = -\,\text{RD}\,.$$

Proof

Proof. This is the result of Solution 0.1. □

Example 0.8 (Switching the reference group in the HERS risk difference)**.** In Example 0.7 the placebo group is the reference group, and the estimated risk difference is 0.007. Taking the hormone therapy group as the reference group instead gives:

```
rd_hat_switched <- risk_hat[["placebo"]] - risk_hat[["hormone therapy"]]
rd_hat_switched
#> [1] -0.0066302
```

This estimate, -0.007, is the negative of 0.007, as Theorem 0.2 states.

Definition 0.7 (Risk ratio (also called the relative risk, relative risk ratio or probability ratio))**.** Let:

- π_1 be the risk of a binary outcome in group 1.
- $\pi_2 > 0$ be the risk of the same outcome in the reference group, group 2.

The **risk ratio**, also called the **relative risk**, **relative risk ratio** or **probability ratio**, of group 1 compared with group 2 is

$$\pi \stackrel{\text{def}}{=} \frac{\pi_1}{\pi_2} \in [0, \infty).$$

Source: Wikipedia contributors [7].

Example 0.9 (Risk ratio of regular exercise in HERS)**.** Take group 1 to be hormone therapy and group 2, the reference group, to be placebo. The estimated risk ratio for the estimated risks of Example 0.5 is:

```
rr_hat <- risk_hat[["hormone therapy"]] / risk_hat[["placebo"]]
rr_hat
#> [1] 1.0173
```

The estimated risk of regular exercise in the hormone therapy group is 1.017 times the estimated risk in the placebo group. This estimate is close to 1, which agrees with the large p-value of Example 0.3: the two groups have about the same estimated risk. The estimate alone does not show how precise it is.

Remark (The risk ratio is not the rate ratio)

Remark 0.5 (The risk ratio is not the rate ratio). Some texts call the risk ratio the rate ratio. The rate ratio is a different quantity: the quotient of two event rates, as in count

regression.

Exercise 0.2. Let:

- π_1 and $\pi_2 > 0$ be the risks of a binary outcome in group 1 and group 2.
- $_\pi \stackrel{\text{def}}{=} \pi_1/\pi_2$ be the risk ratio of group 1 compared with group 2.
- $_\pi > 0$.

Express the risk ratio of group 2 compared with group 1, π_2/π_1, in terms of $_\pi$.

Solution

Solution 0.2.

$$\begin{aligned} \frac{\pi_2}{\pi_1} &= \frac{1}{\pi_1/\pi_2} && \text{(the reciprocal of a quotient swaps numerator and denominator)} \\ &= \frac{1}{_\pi} && \text{(definition of } _\pi\text{)} \end{aligned}$$

Theorem 0.3 (Switching the reference group in a risk ratio)**.** *Let:*

- π_1 *and* $\pi_2 > 0$ *be the risks of a binary outcome in group 1 and group 2.*
- $_\pi \stackrel{def}{=} \pi_1/\pi_2$ *be the risk ratio of group 1 compared with group 2, with group 2 as the reference group.*
- $_\pi > 0$.

Then the risk ratio of group 2 compared with group 1, with group 1 as the reference group, is the reciprocal of $_\pi$*:*

$$\frac{\pi_2}{\pi_1} = \frac{1}{_\pi}.$$

Proof

Proof. This is the result of Solution 0.2. □

Example 0.10 (Switching the reference group in the HERS risk ratio)**.** In Example 0.9 the placebo group is the reference group, and the estimated risk ratio is 1.017. Taking the hormone therapy group as the reference group instead gives:

```
rr_hat_switched <- risk_hat[["placebo"]] / risk_hat[["hormone therapy"]]
rr_hat_switched
#> [1] 0.982993
```

This estimate, 0.983, is the reciprocal of 1.017, 0.983, as Theorem 0.3 states.

Definition 0.8 (Odds ratio)**.** Let:

- ω_1 be the odds of a binary outcome in group 1.
- ω_2 be the odds of the same outcome in the reference group, group 2.

The **odds ratio** of group 1 compared with group 2 is

$$\omega \stackrel{\text{def}}{=} \frac{\omega_1}{\omega_2} \in (0, \infty).$$

Here $\omega_g = \pi_g/(1-\pi_g)$ for the risk π_g of group g.
Source: Wikipedia contributors [10].

Example 0.11 (Odds ratio of regular exercise in HERS)**.** Take group 1 to be hormone therapy and group 2, the reference group, to be placebo. The odds of regular exercise in each group follow from the estimated risks of Example 0.5. The estimated odds ratio divides the odds in the hormone therapy group by the odds in the placebo group:

```
odds_hat <- risk_hat / (1 - risk_hat)
or_hat <- odds_hat[["hormone therapy"]] / odds_hat[["placebo"]]
c(odds_hat, odds_ratio = or_hat)
#>         placebo hormone therapy      odds_ratio
#>        0.621336        0.638955        1.028356
```

The estimated odds ratio 1.028 is close to the risk ratio 1.017 of Example 0.9, because the two groups have nearly equal risks. The odds ratio is 1 exactly when the risk ratio is 1, so it also agrees with the large p-value of Example 0.3.

Exercise 0.3. Let:

- X and Y be binary random variables, each with values 0 and 1.
- $\mathbb{P}(X = x, Y = y) > 0$ for each $x, y \in \{0, 1\}$.

Express the odds of $Y = 1$ given $X = x$ in terms of the joint probabilities $\mathbb{P}(X = x, Y = 1)$ and $\mathbb{P}(X = x, Y = 0)$.

Solution

Solution 0.3.

$$\begin{aligned}
\omega &= \frac{\mathbb{P}(Y=1 \mid X=x)}{1-\mathbb{P}(Y=1 \mid X=x)} && \text{(definition of odds)} \\
&= \frac{\mathbb{P}(Y=1 \mid X=x)}{\mathbb{P}(Y=0 \mid X=x)} && \text{(the two values of } Y \text{ have conditional probabilities that sum to 1)} \\
&= \frac{\mathbb{P}(X=x, Y=1)/\mathbb{P}(X=x)}{\mathbb{P}(X=x, Y=0)/\mathbb{P}(X=x)} && \text{(definition of conditional probability, in numerator and denominator)} \\
&= \frac{\mathbb{P}(X=x, Y=1)}{\mathbb{P}(X=x, Y=0)} && \text{(divide the numerator and the denominator by the same positive number}
\end{aligned}$$

Theorem 0.4 (The odds ratio as a ratio of cross products)**.** *Let:*

- X *and* Y *be binary random variables, each with values* 0 *and* 1*.*
- $\mathbb{P}(X=x, Y=y) > 0$ *for each* $x, y \in \{0,1\}$*.*

The odds ratio of $Y=1$ *comparing the group* $X=1$ *with the reference group* $X=0$ *is*

$$_{\omega}(Y \mid X) = \frac{\mathbb{P}(X=1, Y=1)\mathbb{P}(X=0, Y=0)}{\mathbb{P}(X=1, Y=0)\mathbb{P}(X=0, Y=1)}.$$

Proof

Proof. In the group $X=x$, the odds of $Y=1$ are $\mathbb{P}(Y=1 \mid X=x)/\mathbb{P}(Y=0 \mid X=x)$, because the two conditional probabilities sum to 1. Writing each conditional probability as a joint probability over $\mathbb{P}(X=x)$, which is nonzero, gives

$$\frac{\mathbb{P}(Y=1 \mid X=x)}{\mathbb{P}(Y=0 \mid X=x)} = \frac{\mathbb{P}(X=x, Y=1)/\mathbb{P}(X=x)}{\mathbb{P}(X=x, Y=0)/\mathbb{P}(X=x)} = \frac{\mathbb{P}(X=x, Y=1)}{\mathbb{P}(X=x, Y=0)},$$

where the last step multiplies the numerator and denominator by the same nonzero factor $\mathbb{P}(X=x)$. So the odds of $Y=1$ in the group $X=1$ are $\mathbb{P}(X=1, Y=1)/\mathbb{P}(X=1, Y=0)$, and in the group $X=0$ the odds are $\mathbb{P}(X=0, Y=1)/\mathbb{P}(X=0, Y=0)$. The odds ratio is the first odds divided by the second. Dividing by a fraction is multiplying by its reciprocal, which gives the stated product of the four joint probabilities. □

Example 0.12 (The HERS odds ratio from the cell proportions)**.** Take $X=1$ for hormone therapy and $X=0$ for placebo, and $Y=1$ for regular exercise. The estimated

joint probabilities are the cell proportions of the HERS table:

```
joint_hat <- prop.table(table(HT = hers$HT, exercise = hers$exercise))
cross_ratio <-
  joint_hat["hormone therapy", "yes"] * joint_hat["placebo", "no"] /
  (joint_hat["hormone therapy", "no"] * joint_hat["placebo", "yes"])
cross_ratio
#> [1] 1.02836
```

The cross-product expression equals 1.028. It equals the estimated odds ratio 1.028 of Example 0.11, as Theorem 0.4 states.

Theorem 0.5 (The odds ratio is reversible)**.** *Let:*

- X *and* Y *be binary random variables, each with values* 0 *and* 1*.*
- $\mathbb{P}(X = x, Y = y) > 0$ *for each* $x, y \in \{0, 1\}$*.*

The odds ratio of $Y = 1$ *comparing* $X = 1$ *with* $X = 0$ *equals the odds ratio of* $X = 1$ *comparing* $Y = 1$ *with* $Y = 0$*:*

$$_{\omega}(Y \mid X) = \ _{\omega}(X \mid Y).$$

Proof

Proof. By Theorem 0.4, with the roles of X and Y exchanged,

$$_{\omega}(X \mid Y) = \frac{\mathbb{P}(Y = 1, X = 1)\mathbb{P}(Y = 0, X = 0)}{\mathbb{P}(Y = 1, X = 0)\mathbb{P}(Y = 0, X = 1)}.$$

The event $\{Y = y, X = x\}$ is the event $\{X = x, Y = y\}$, so each joint probability here equals the joint probability with its arguments in the other order. The denominator is a product of two joint probabilities, and multiplication is commutative. Reordering its two factors gives the denominator of Theorem 0.4 for $_{\omega}(Y \mid X)$. □

Example 0.13 (Reversing the roles of treatment and exercise in HERS)**.** In Example 0.12, the odds ratio compares the odds of regular exercise between the hormone therapy and placebo groups. Reversing the roles compares the odds of hormone therapy between the women who exercise regularly and those who do not:

```r
counts <- table(HT = hers$HT, exercise = hers$exercise)
odds_ht_exercise <- counts["hormone therapy", "yes"] / counts["placebo", "yes"]
odds_ht_no_exercise <- counts["hormone therapy", "no"] / counts["placebo", "no"]
or_reversed <- odds_ht_exercise / odds_ht_no_exercise
or_reversed
#> [1] 1.02836
```

The reversed odds ratio is 1.028. It equals the odds ratio 1.028 of Example 0.11, as Theorem 0.5 states.

Theorem 0.6 (The conditional odds ratio is reversible)**.** *Let:*

- X *and* Y *be binary random variables, each with values* 0 *and* 1*.*
- Z *be a random variable, and* z *be a value of* Z*.*
- $\mathbb{P}(X = x, Y = y, Z = z) > 0$ *for each* $x, y \in \{0, 1\}$*.*

The odds ratio of $Y = 1$ *comparing* $X = 1$ *with the reference group* $X = 0$*, among the units with* $Z = z$*, equals the odds ratio of* $X = 1$ *comparing* $Y = 1$ *with* $Y = 0$*, among the units with* $Z = z$*:*

$$\text{OR}(Y \mid X, Z = z) = \text{OR}(X \mid Y, Z = z).$$

Proof

Proof. First, find the odds of $Y = 1$ in the group $X = x$ among the units with $Z = z$:

$$\begin{aligned}
\omega_x &= \frac{\mathbb{P}(Y = 1 \mid X = x, Z = z)}{1 - \mathbb{P}(Y = 1 \mid X = x, Z = z)} && \text{(definition of odds)} \\
&= \frac{\mathbb{P}(Y = 1 \mid X = x, Z = z)}{\mathbb{P}(Y = 0 \mid X = x, Z = z)} && \text{(the two values of } Y \text{ have conditional probabilities that sur} \\
&= \frac{\mathbb{P}(X = x, Y = 1, Z = z)/\mathbb{P}(X = x, Z = z)}{\mathbb{P}(X = x, Y = 0, Z = z)/\mathbb{P}(X = x, Z = z)} && \text{(definition of conditional probability, in numerator and den} \\
&= \frac{\mathbb{P}(X = x, Y = 1, Z = z)}{\mathbb{P}(X = x, Y = 0, Z = z)} && \text{(divide the numerator and the denominator by the same po}
\end{aligned}$$

The odds ratio is the odds ω_1 divided by the odds ω_0. Dividing by a fraction is multiplying by its reciprocal, so

$$\text{OR}(Y \mid X, Z = z) = \frac{\mathbb{P}(X = 1, Y = 1, Z = z)\mathbb{P}(X = 0, Y = 0, Z = z)}{\mathbb{P}(X = 1, Y = 0, Z = z)\mathbb{P}(X = 0, Y = 1, Z = z)}.$$

The same steps with the roles of X and Y exchanged give

$$_{\omega}(X \mid Y, Z = z) = \frac{\mathbb{P}(Y = 1, X = 1, Z = z)\mathbb{P}(Y = 0, X = 0, Z = z)}{\mathbb{P}(Y = 1, X = 0, Z = z)\mathbb{P}(Y = 0, X = 1, Z = z)}.$$

The event $\{Y = y, X = x, Z = z\}$ is the event $\{X = x, Y = y, Z = z\}$, so each joint probability here equals the joint probability with its arguments in the other order. The denominator is a product of two joint probabilities, and multiplication is commutative. Reordering its two factors gives the denominator of the first display. □

Example 0.14 (Reversing the roles of treatment and exercise within diabetes strata in HERS)**.** In Example 0.13, the roles of hormone therapy and regular exercise are reversed among all the women. Here the same two odds ratios are computed separately for the women with diabetes and the women without it:

```
or_by_diabetes <- sapply(c("no", "yes"), function(z) {
  counts <- table(
    HT = hers$HT[hers$diabetes == z],
    exercise = hers$exercise[hers$diabetes == z]
  )
  ht <- "hormone therapy"
  c(
    ht_odds_ratio = unname(
      (counts[ht, "yes"] / counts["placebo", "yes"]) /
        (counts[ht, "no"] / counts["placebo", "no"])
    ),
    exercise_odds_ratio = unname(
      (counts[ht, "yes"] / counts[ht, "no"]) /
        (counts["placebo", "yes"] / counts["placebo", "no"])
    )
  )
})
or_by_diabetes
#>                          no     yes
#> ht_odds_ratio       1.02222 1.08842
#> exercise_odds_ratio 1.02222 1.08842
```

The first row compares the odds of hormone therapy between the women who exercise regularly and those who do not. The second row compares the odds of regular exercise between the hormone therapy and placebo groups. Among the women without diabetes the two odds ratios are 1.022 and 1.022, and among the women with diabetes they are 1.088 and 1.088. In each stratum the two agree, as Theorem 0.6 states. The stratum odds ratios differ from the odds ratio 1.028 of Example 0.11 for all the women, because the theorem equates the two odds ratios within a stratum, not either one with the marginal

odds ratio.

Definition 0.9 (Case)**.** In a study that selects its participants by whether they have an outcome of interest, a **case** is a participant who has the outcome [8, Ch. 8].

Definition 0.10 (Control)**.** In a study that selects its participants by whether they have an outcome of interest, a **control** is a participant who does not have the outcome, sampled from the same population as the cases [8, Ch. 8].

Definition 0.11 (Case-control study)**.** A **case-control study** is an observational study that selects its participants by outcome: it includes the cases of a population and a sample of controls from the same population, and then compares the exposure of the two groups [8, Ch. 8].

Example 0.15 (A case-control study of regular exercise in HERS)**.** Take the women who exercise regularly as the cases. Draw as many controls as there are cases, at random, from the women who do not exercise regularly. Then compare the odds of hormone therapy between the cases and the controls:

```
set.seed(571)
cases <- hers[hers$exercise == "yes", ]
non_cases <- hers[hers$exercise == "no", ]
controls <- non_cases[sample(nrow(non_cases), nrow(cases)), ]
odds_ht <- function(d) {
  sum(d$HT == "hormone therapy") / sum(d$HT == "placebo")
}
or_case_control <- odds_ht(cases) / odds_ht(controls)
or_case_control
#> [1] 1.03816
```

The study has 1068 cases and 1068 controls. Its odds ratio is 1.038, near the odds ratio 1.028 of Example 0.11 from the whole cohort. The two are close because the odds ratio is reversible (Theorem 0.5): sampling by outcome does not change it, although it would change the risks.

> **Remark** (Controls in experiments)
>
> *Remark* 0.6 (Controls in experiments). The word "control" has a second meaning. In an experiment, the investigator assigns the exposure. The **controls** are the participants who receive the baseline treatment for comparison, such as a placebo [8, Ch. 6]. These controls are chosen by assignment, not by outcome. They are not the controls of a case-control study.

> **Read**
>
> Rice [5, Sec. 13.6], pp. 526–529 covers odds ratios.

Further reading

- Sections 4.6 to 4.8 of Lange's *Mathematical and Statistical Methods for Genetic Analysis* [11] cover exact tests of independence, case-control association tests, and the transmission/disequilibrium test, with genetic examples.

References

[1] E. Vittinghoff, D. V. Glidden, S. C. Shiboski, and C. E. McCulloch, *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models*, 2nd ed. Springer, 2012. doi: 10.1007/978-1-4614-1353-0.

[2] R. V. Hogg, E. A. Tanis, and D. L. Zimmerman, *Probability and statistical inference*, Tenth edition. Boston: Pearson, 2019.

[3] Wikipedia contributors, "Contingency table — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Contingency_table

[4] Wikipedia contributors, "Pearson's chi-squared test — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Pearson%27s_chi-squared_test

[5] J. A. Rice, *Mathematical statistics and data analysis*, 3rd ed. Belmont, CA: Cengage Learning, 2007.

[6] Wikipedia contributors, "Fisher's exact test — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Fisher%27s_exact_test

[7] Wikipedia contributors, "Relative risk — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Relative_risk

[8] K. J. Rothman, T. L. Lash, T. J. VanderWeele, and S. Haneuse, *Modern epidemiology*, 4th ed. Philadelphia: Wolters Kluwer, 2021.

[9] Wikipedia contributors, "Absolute risk reduction — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Absolute_risk_reduction

[10] Wikipedia contributors, "Odds ratio — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Odds_ratio

[11] K. Lange, *Mathematical and statistical methods for genetic analysis*, 2nd ed. in Statistics for biology and health. Springer, 2002. doi: 10.1007/978-0-387-21750-5.