



[← Previous](#)



[Next >](#)

Questions 3 and 4: Esophageal cancer and alcohol/tobacco use, part 1

[Bookmark this page](#)

Assessment due Mar 5, 2021 23:52 IST

Case-control studies help determine whether certain exposures are associated with outcomes such as developing cancer. The built-in dataset `esoph` contains data from a case-control study in France comparing people with esophageal cancer (cases, counted in `ncases`) to people without esophageal cancer (controls, counted in `ncontrols`) that are carefully matched on a variety of demographic and medical characteristics. The study compares alcohol intake in grams per day (`alcgp`) and tobacco intake in grams per day (`tobgp`) across cases and controls grouped by age range (`agegp`).

The dataset is available in base R and can be called with the variable name `esoph`:

```
head(esoph)
```

You will be using this dataset to answer the following four multi-part questions (Questions 3-6).

You may wish to use the `tidyverse` package:

```
library(tidyverse)
```

The following three parts have you explore some basic characteristics of the dataset.

Each row contains one group of the experiment. Each group has a different combination of age, alcohol consumption, and tobacco consumption. The number of cancer cases and number of controls (individuals without cancer) are reported for each group.

Question 3a

1.0/1.0 point (graded)

How many groups are in the study?



88

Submit

You have used 1 of 10 attempts

Question 3b

1.0/1.0 point (graded)

How many cases are there?

Save this value as `ncases` for later problems

Save this value as `all_cases` for later problems.

✓

200

Submit

You have used 1 of 10 attempts

Question 3c

1.0/1.0 point (graded)

How many controls are there?

Save this value as `all_controls` for later problems.

✓

975

Submit

You have used 1 of 10 attempts

The following four parts ask you to explore some probabilities within this dataset related to alcohol and tobacco consumption.

Question 4a

0.0/1.0 point (graded)

What is the probability that a subject in the highest alcohol consumption group is a cancer case?

Submit

You have used 0 of 10 attempts

Question 4b

0.0/1.0 point (graded)

What is the probability that a subject in the lowest alcohol consumption group is a cancer case?

Submit

You have used 0 of 10 attempts

Question 4c

1.0/1.0 point (graded)

Given that a person is a case, what is the probability that they smoke 10g or more a day?

✓

0.61

Question 4d

1.0/1.0 point (graded)
Given that a person is a control, what is the probability that they smoke 10g or more a day?

0.4615385

✓

0.4615385



edX

- [About](#)
- [Affiliates](#)
- [edX for Business](#)
- [Open edX](#)
- [Careers](#)
- [News](#)

Legal

- [Terms of Service & Honor Code](#)
- [Privacy Policy](#)

[Privacy Policy.](#)
[Accessibility Policy.](#)
[Trademark Policy.](#)
[Sitemap](#)

Connect

[Blog](#)
[Contact Us](#)
[Help Center](#)
[Media Kit](#)
[Donate](#)



© 2021 edX Inc. All rights reserved.
深圳市恒宇博科技有限公司 [粤ICP备17044299号-2](#)