ST552 Statistical Methods II

Permutation Tests

Feb 12 2016

Charlotte Wickham

stat552.cwick.co.nz

Randomization test

The F-tests rely on the Normal error assumption.

In a randomized experiment, the randomization provides a basis for inference (no i.i.d sampling from populations required) and results in **randomization** tests.

The same procedure can be used in observational studies, with the assumption that nature ran the experiment for you, i.e. it's like units were assigned to values of the explanatory variables at random.

Some people call randomization tests used for observational data, **permutation** tests.

Your turn

Group exercise:

What are the key ingredients in a hypothesis test?

How do you find a p-value?

Example: Overall regression F-test

Model: $Y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I)$,

X_{nxp} fixed, $\beta = (\beta_0\ \beta_1\ \ldots\ \beta_{p-1})$ unknown

Null hypothesis: $\beta_1 = \ldots = \beta_{p-1} = 0$

Test statistic:

$$F = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2 - \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{(p-1)} \div \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-p}$$

Null distribution: F-distribution with p-1 and n-p degrees of freedom.

Randomization test

Model: Randomized experiment

Null hypothesis: treatment has no effect on response.

Test statistic:

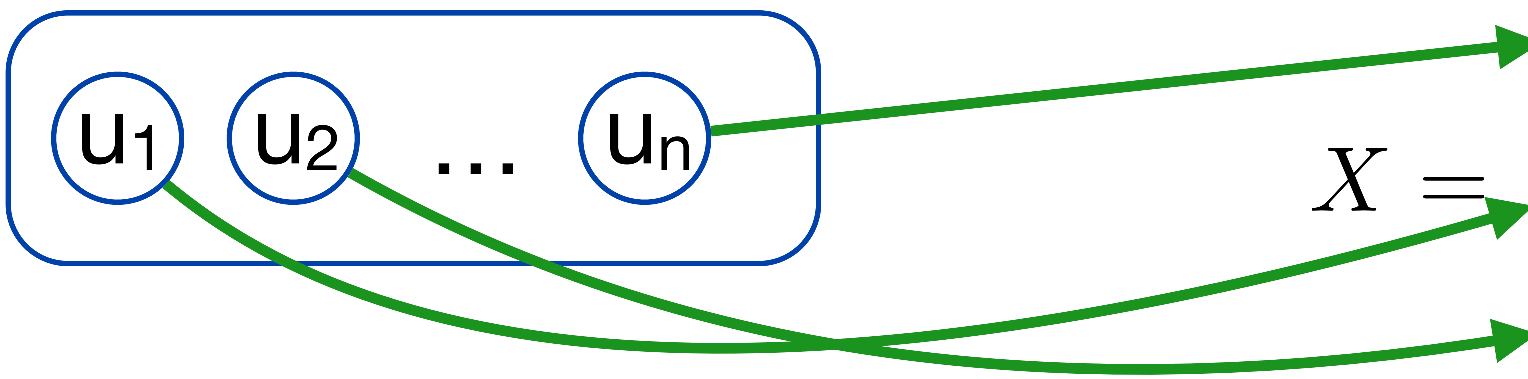
Should quantify evidence against the null.

Let’s use the usual overall regression F-stat.

Null distribution: Randomization distribution of test-statistic

The randomized experiment model

n experimental units



Fixed X

$$X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1(p-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{n(p-1)} \end{pmatrix}$$

describes treatments

randomly assign units to treatments

every permutation of units to treatments is equally likely

apply treatments, observe response

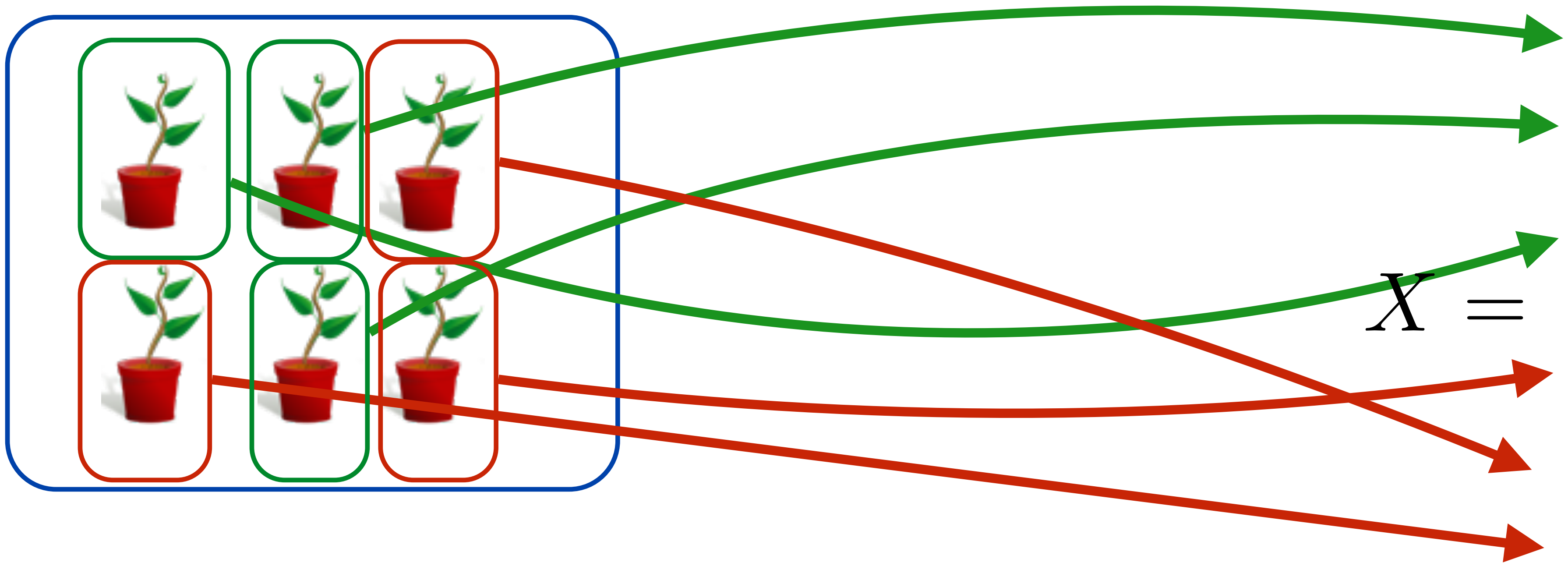
$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

y_i is the observed response for the unit that was randomly assigned to the ith row of the design matrix

The randomized experiment model: growing tomatoes

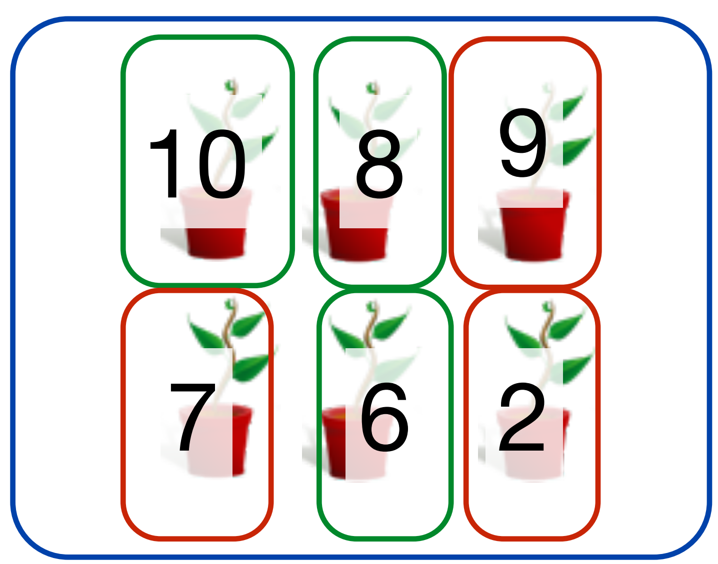
n experimental units

received fertilizer



$$X = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$$

apply treatments, observe response



$$y = \begin{pmatrix} 8 \\ 6 \\ 10 \\ 2 \\ 9 \\ 7 \end{pmatrix}$$

Null distribution

If the null is true: treatments have no effect on response.

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

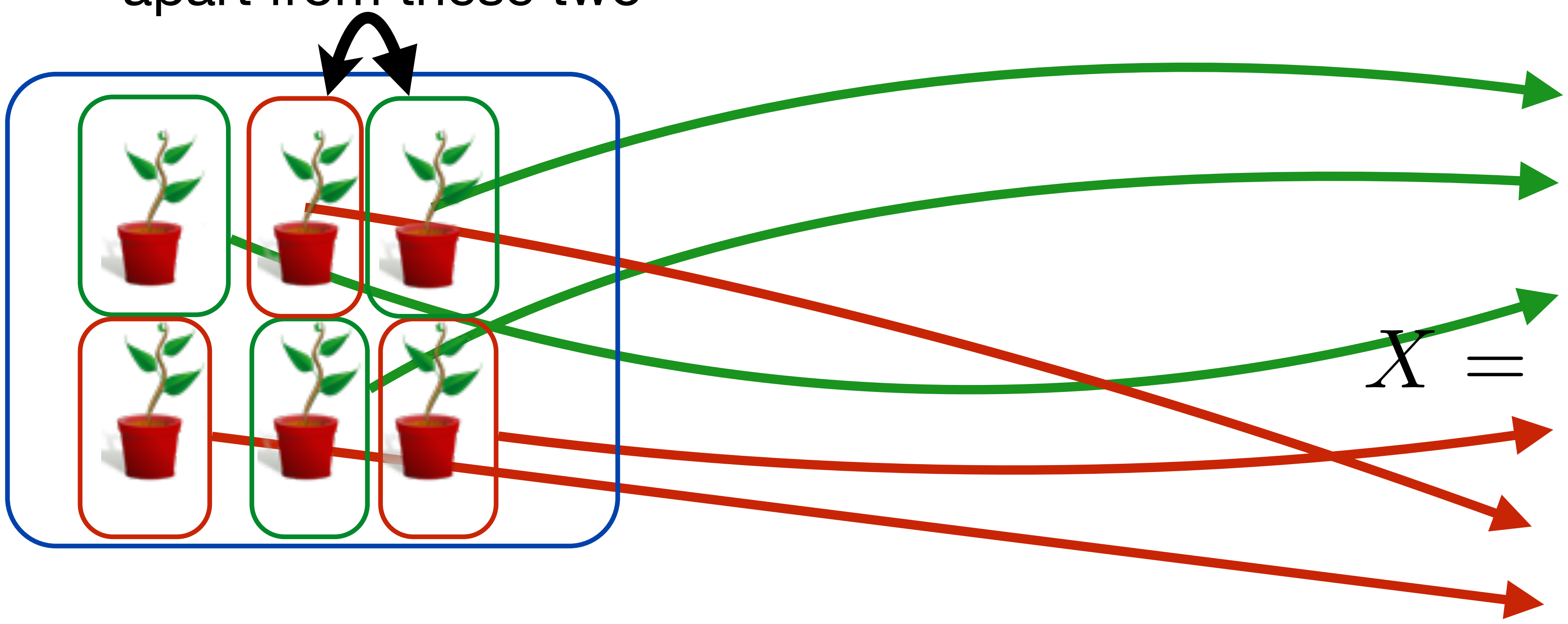
y_i is the observed response for the unit that was randomly assigned to the ith row of the design matrix

If the null is true, I see the same set of y_i, just in different order based on the output of my randomizing units to treatment.

Null distribution: the distribution of the test-statistic for all permutations of y_i

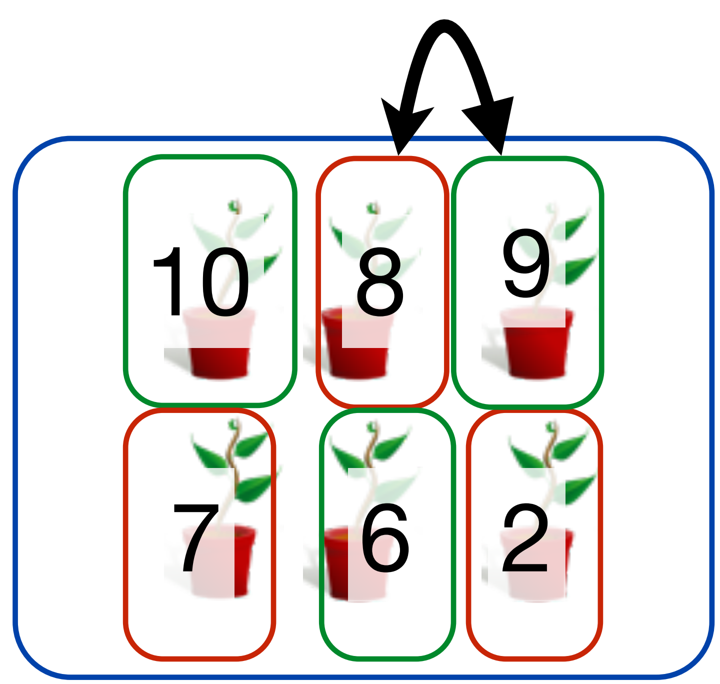
An equally likely outcome under the null hypothesis: growing tomatoes

this randomization is identical
apart from these two



$$X = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$$

apply treatments,
observe response



same responses, since
treatment has no effect

$$y = \begin{pmatrix} 9 \\ 6 \\ 10 \\ 2 \\ 8 \\ 7 \end{pmatrix}$$

permuted
order

Observed data:

$$y = \begin{pmatrix} 8 \\ 6 \\ 10 \\ 2 \\ 9 \\ 7 \end{pmatrix} \quad X = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$$

Observed F-stat = 17.6

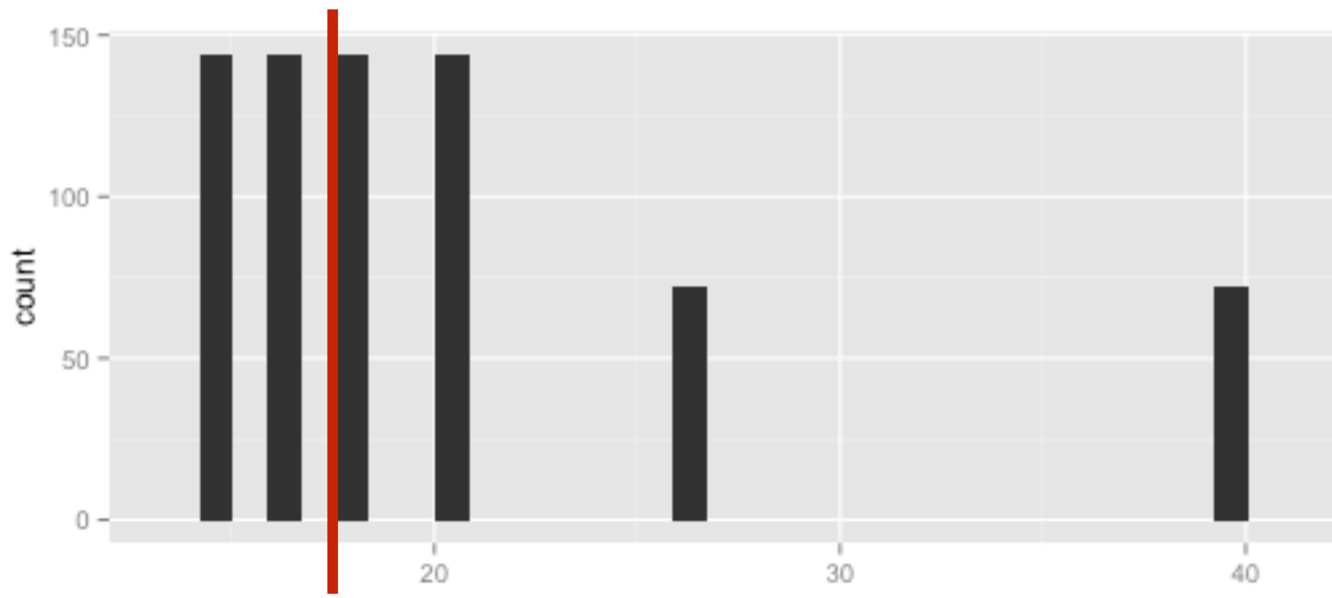
Other equally likely y's under the null

2	7	9	2	6	9	10	9	2
10	10	8	6	10	7	7	7	7
6	9	7	9	2	2	2	10	10
8	8	6	8	8	6	8	8	6
7	2	2	10	7	10	9	2	8
9	6	10	7	9	8	6	6	9

... 711 other permutations

Other equally F-stats under the null

17.6 26.6 17.6 20.8 17.6 17.6 15.9 26.6 15.9 ... 711 other F-stats



Null distribution

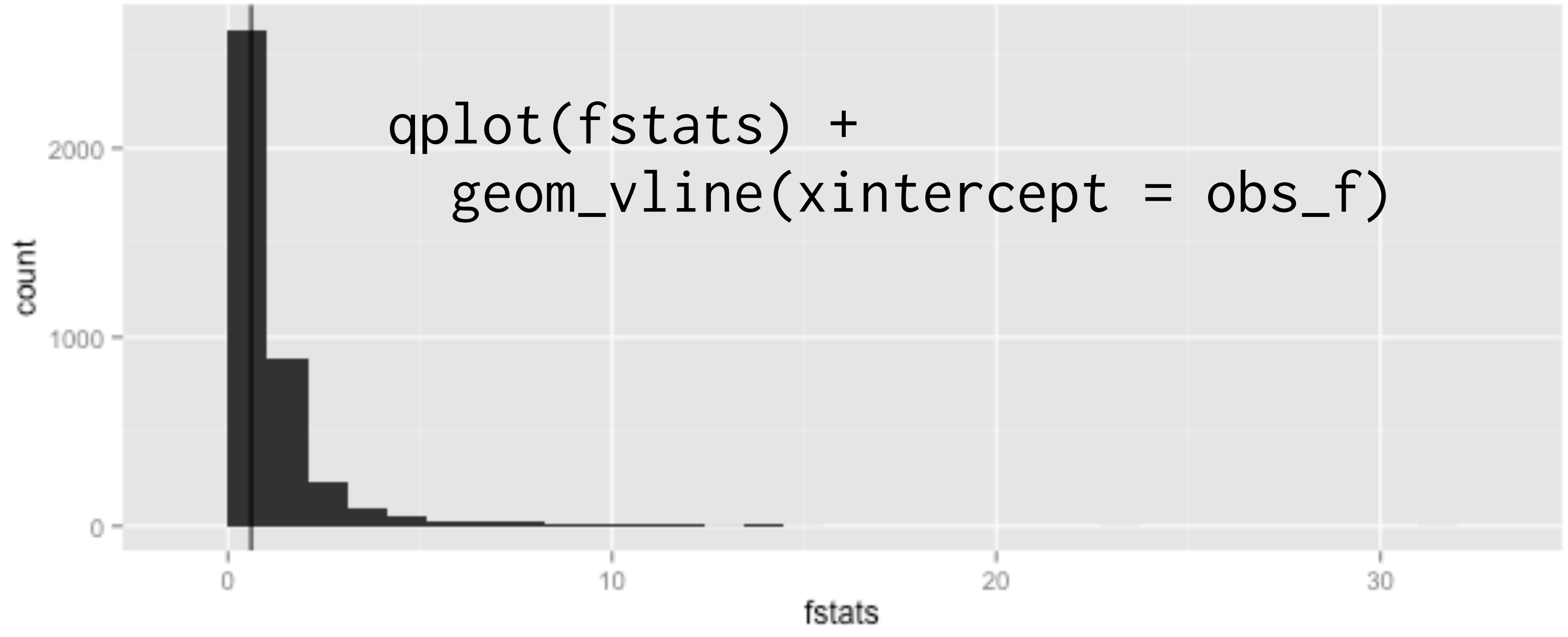
Faraway: Galapagos

```
library(faraway)
data(gala, package="faraway")
lmod <- lm(Species ~ Nearest + Scruz, gala)

# observed overall regression F-stat
obs_f <- summary(lmod)$fstatistic["value"]

nreps <- 4000
set.seed(123)
fstats <- numeric(nreps)

for(i in 1:nreps){
  lmods <- lm(sample(Species) ~ Nearest + Scruz, gala)
  fstats[i] <- summary(lmods)$fstatistic["value"]
}
```




```
qplot(fstats) +
  geom_vline(xintercept = obs_f)
```

Inference in complete populations

Faraway makes a comment “permutation tests give some meaning to the p-value for the sample at hand”

What does he mean?

We observed the complete population, but for each unit we only observed it’s response at one set of explanatory values.

That is, we observed only one of many populations that could have been generated by reassigning units different explanatory values.

The p-value from the permutation/randomization test quantifies how likely that data was under the hypothesis the explanatory variables have no effect, assuming it was like the units were assigned to explanatory values at random.