

Stat310

Hypothesis testing

Hadley Wickham

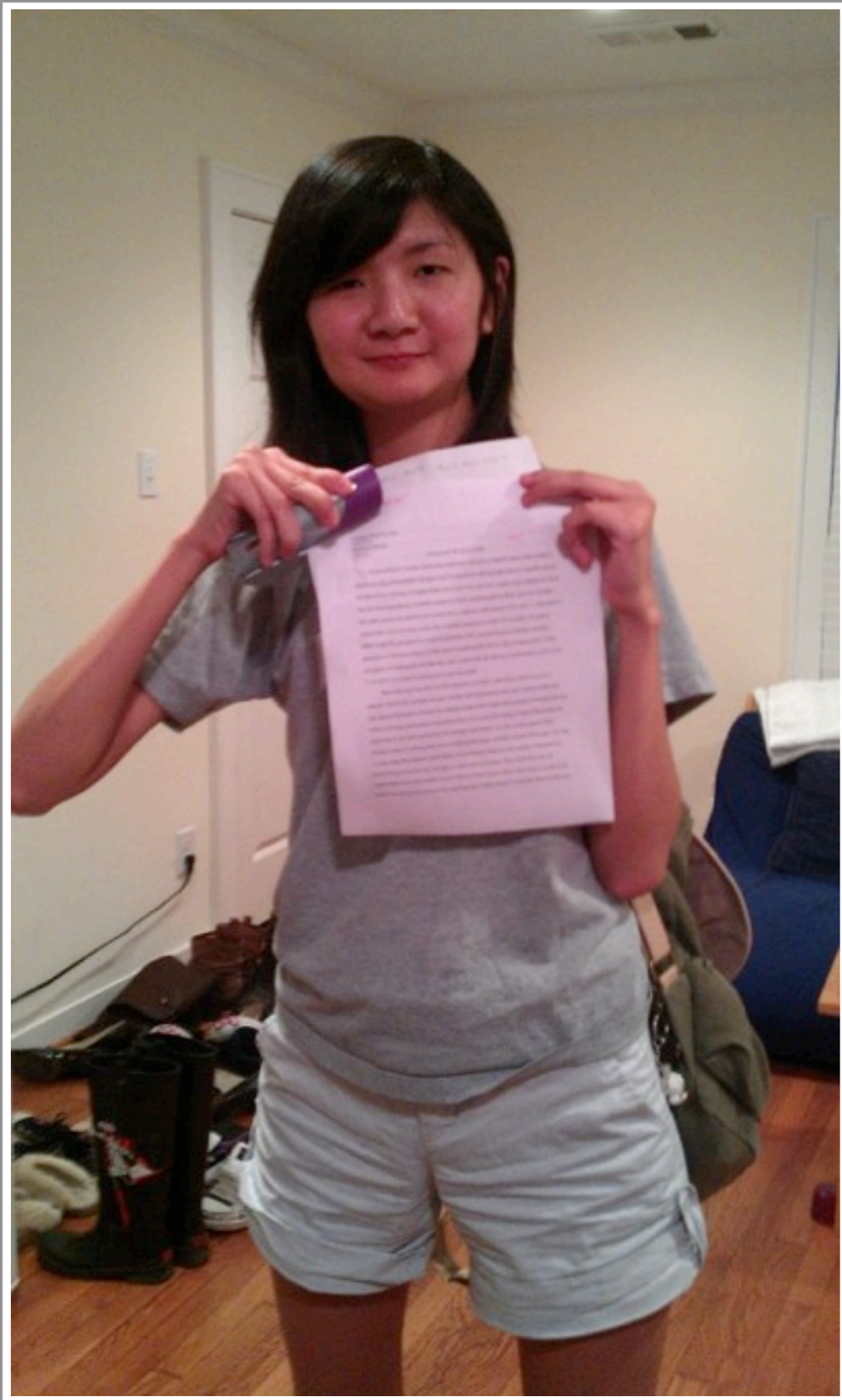
1. Test & Final

2. Review

3. Hypothesis testing

4. Case study

5. t-test



Test

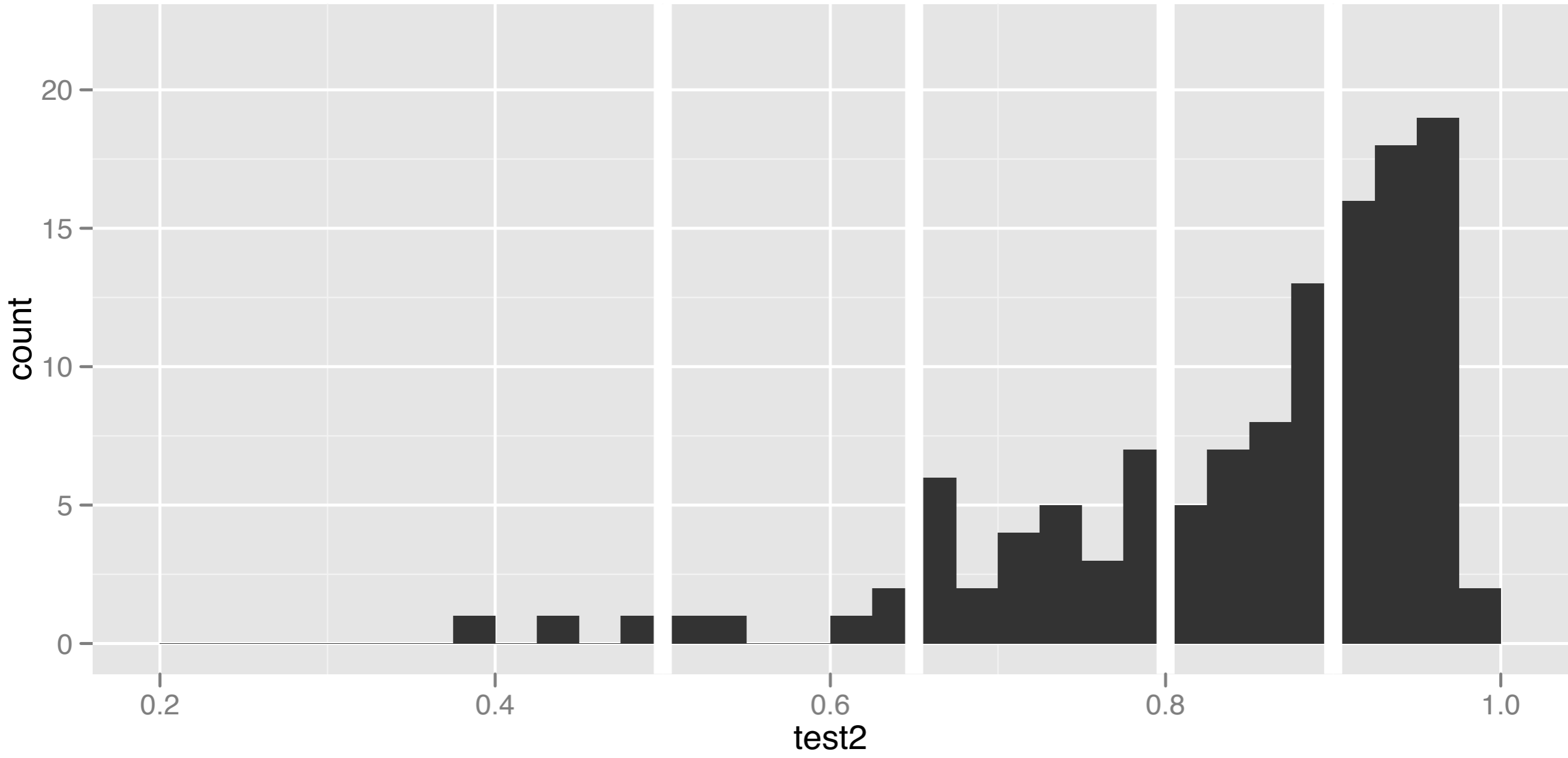
F

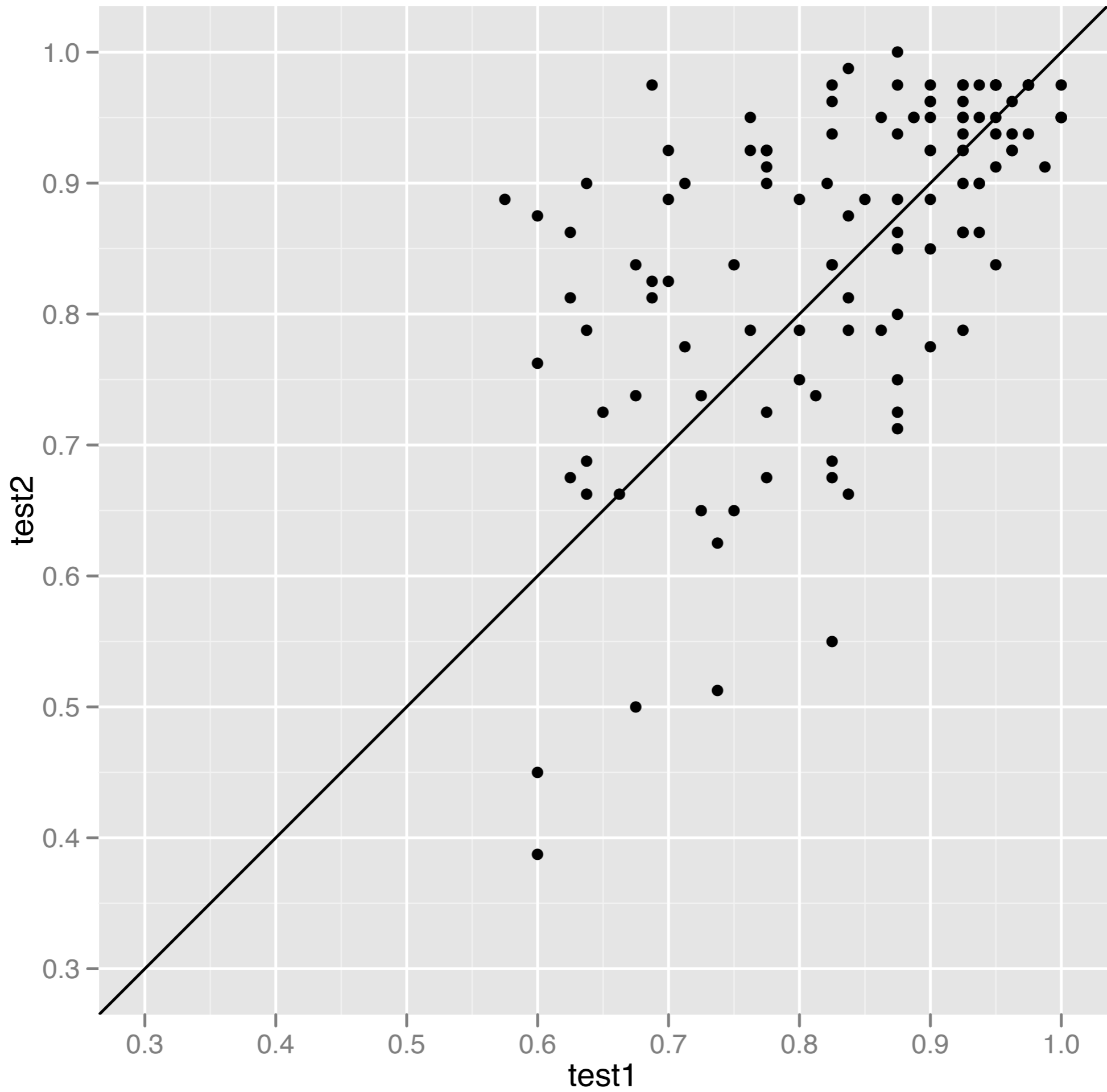
D

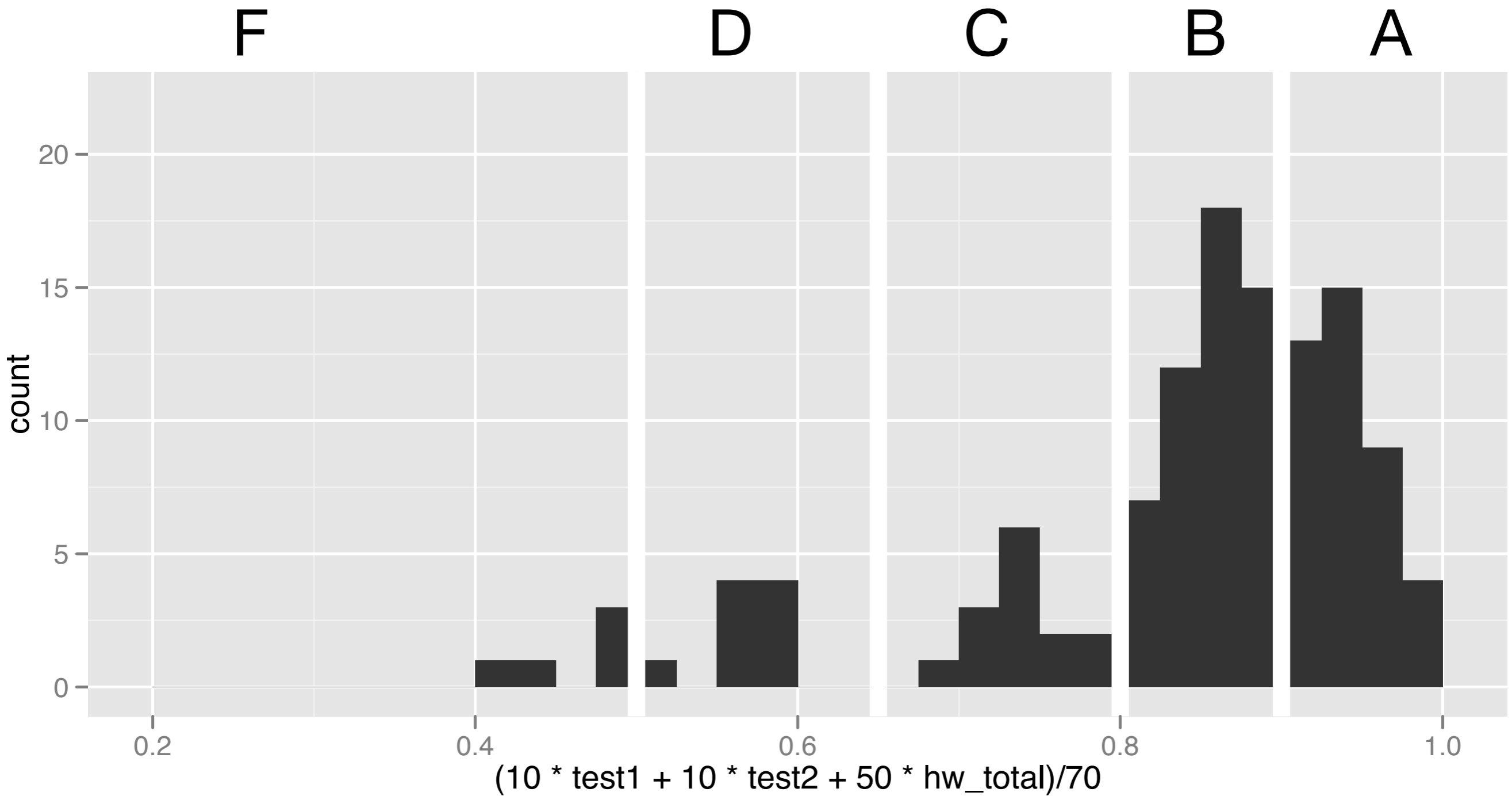
C

B

A







Doesn't account for dropping lowest two.
 Does account for bonus homeworks.

Questions

1 was a bit tricky (because I forgot that it was a complex integral), but you weren't penalised if you missed it.

2 required a bit of thought. And "relaxing assumptions" was a bit tricky.

3 was straightforward. But remember the component of and is or

4 should have also been straightforward

Final

Final details

2 hour take home cumulative final.

Available Wed April 25.

Due Wed May 4 @ 5pm.

Three pages (six sides) of notes.

4 questions: two theoretical, two applied.

Stats in practice

One more, due end of finals week.

Will have better guide up before then.

Bonus: stats in real life

Take a picture or video that raises a question that you could solve with your stat310 skills.

Inspiration:

<http://www.101qs.com/top10.php>, <http://www.statpics.com/>

Reviewing

Next week, we will spend some time reviewing material.

Will have study help sessions before/during finals week.

Post your questions here:

<http://www.google.com/moderator/#15/e=1fe695&t=1fe695.40>

Review

Steps

Identify distribution that connects estimator and true value.

Form confidence interval for known (sampling) distribution.

Write as probability statement.

Back transform.

Write as interval.

So far

Given X_i iid $\text{SomeDist}(a, b)$, and **data**,
under what circumstances can we give
estimates of the uncertainty of
parameters a and b ?

Next: We'll learn how to answer questions
like is $a > 0$, or is b in $[5, 10]$.

Variance

X_i iid Normal(5, σ^2), $i = 1, \dots, 10$

I ran the experiment and recorded the following results: 0.15 1.48 1.25 2.47 1.09 1.95 1.46 1.49 2.81 1.96 ($s^2 = 0.55$)

Find a 95% confidence interval for the variance.
(Hint: If $X \sim \chi^2(9)$, $F_X(19) = 0.975$ and $F_X(2.7) = 0.025$).

EC: Can you also make a confidence interval for the standard deviation?

Your turn

The following values have generated from iid Normal(μ , 1):

2.9 2.1 3.0 3.2 1.2 3.0 3.3 1.2 2.3 1.5
(mean: 2.37)

Is it plausible that they came from a normal distribution with mean 1.5?

Hypothesis testing

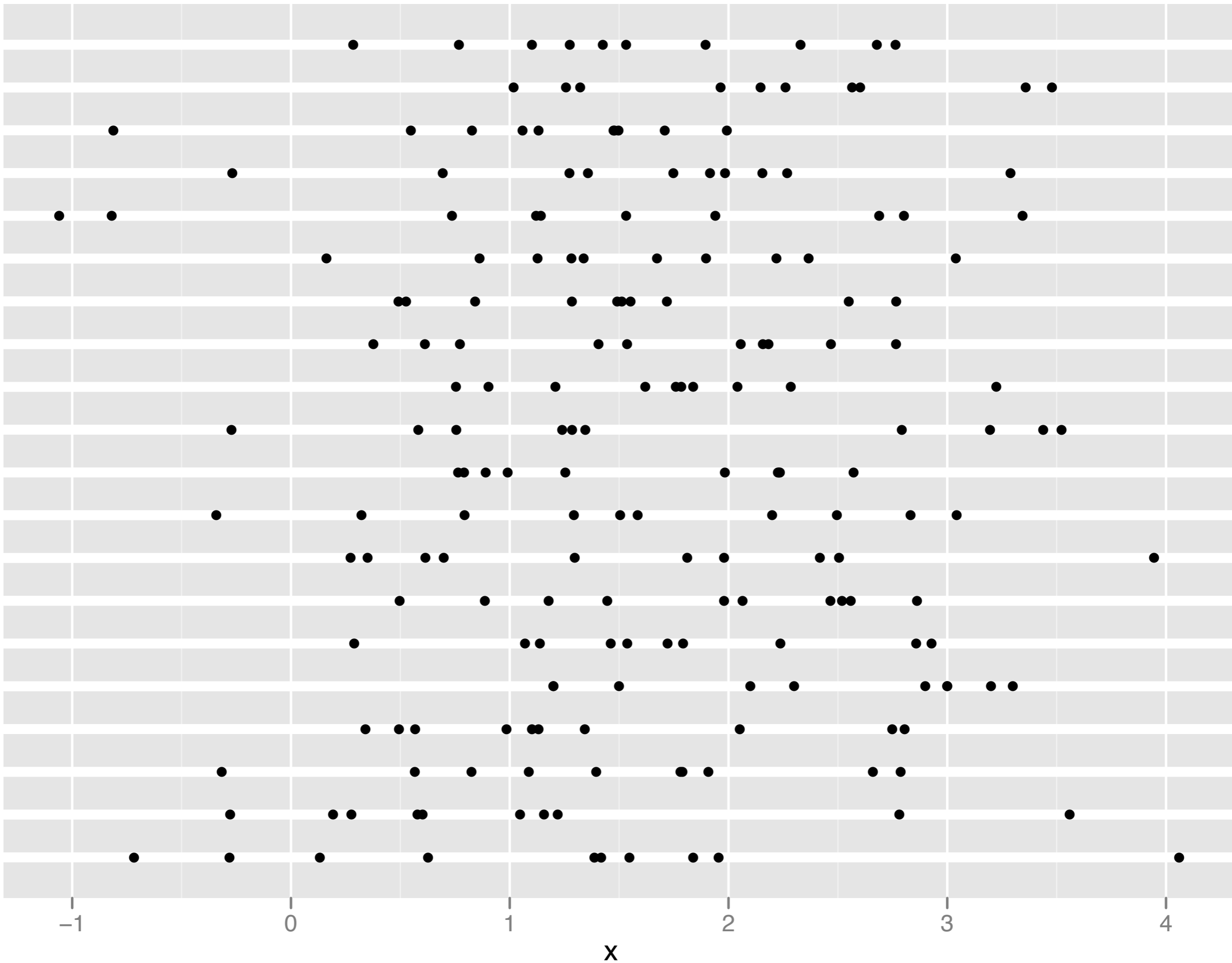
The statistical justice system

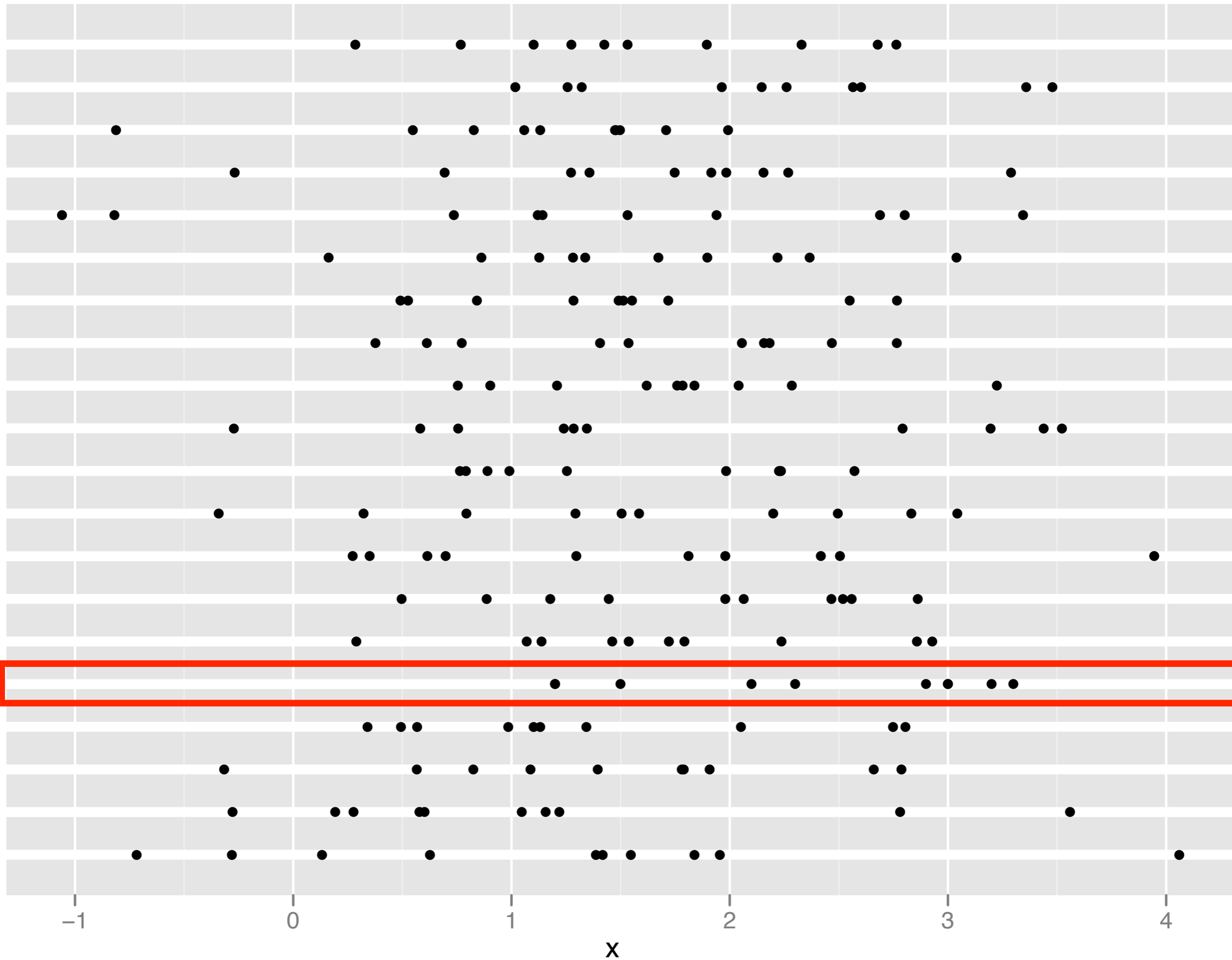


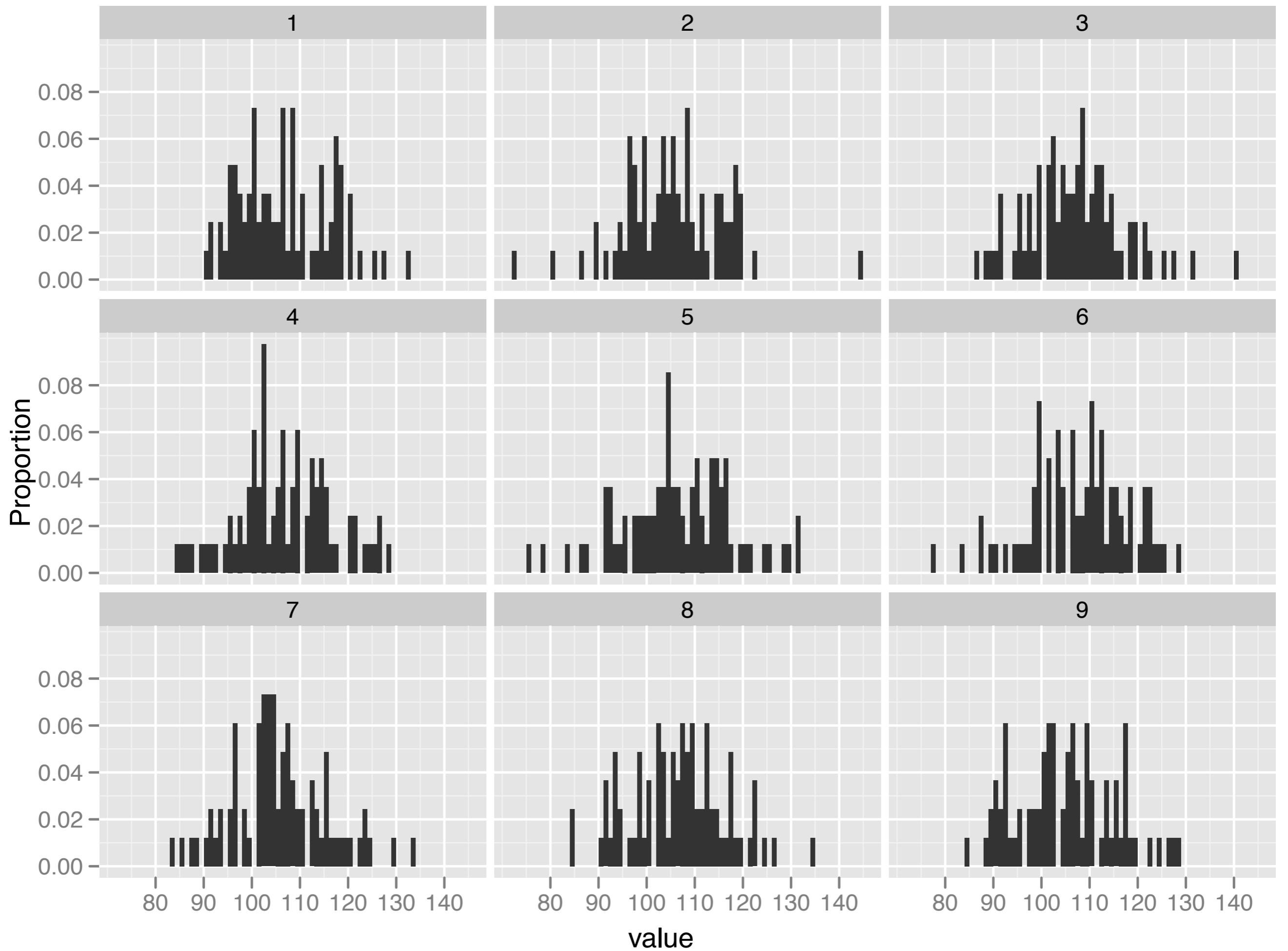
A **suspect** is accused of a **crime**. The suspect is declared guilty or not guilty based on a **trial**. Each trial has a **defence** and a **prosecution**. On the basis of how **evidence** compares to a **standard**, the judge makes a decision to **convict** or **acquit**.

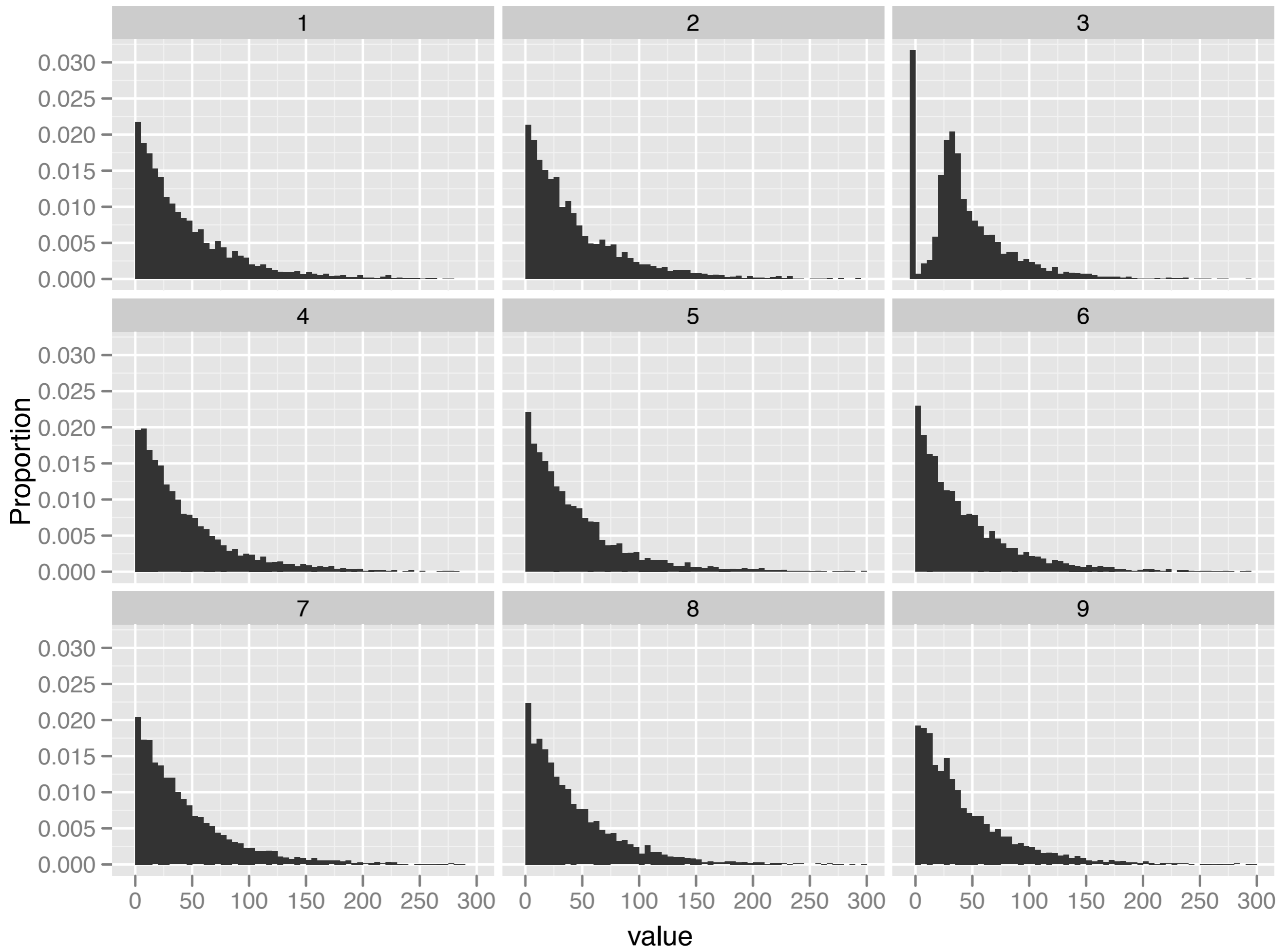
A **distribution** is accused of not having a **particular parameter**. The data is declared guilty or not guilty based on the results of a **statistical test**. Each test has a **null hypothesis** and an **alternative hypothesis**. On the basis of how a **test statistic** compares to a standard **distribution**, we make the decision to **reject the null** or **fail to reject the null hypothesis**.

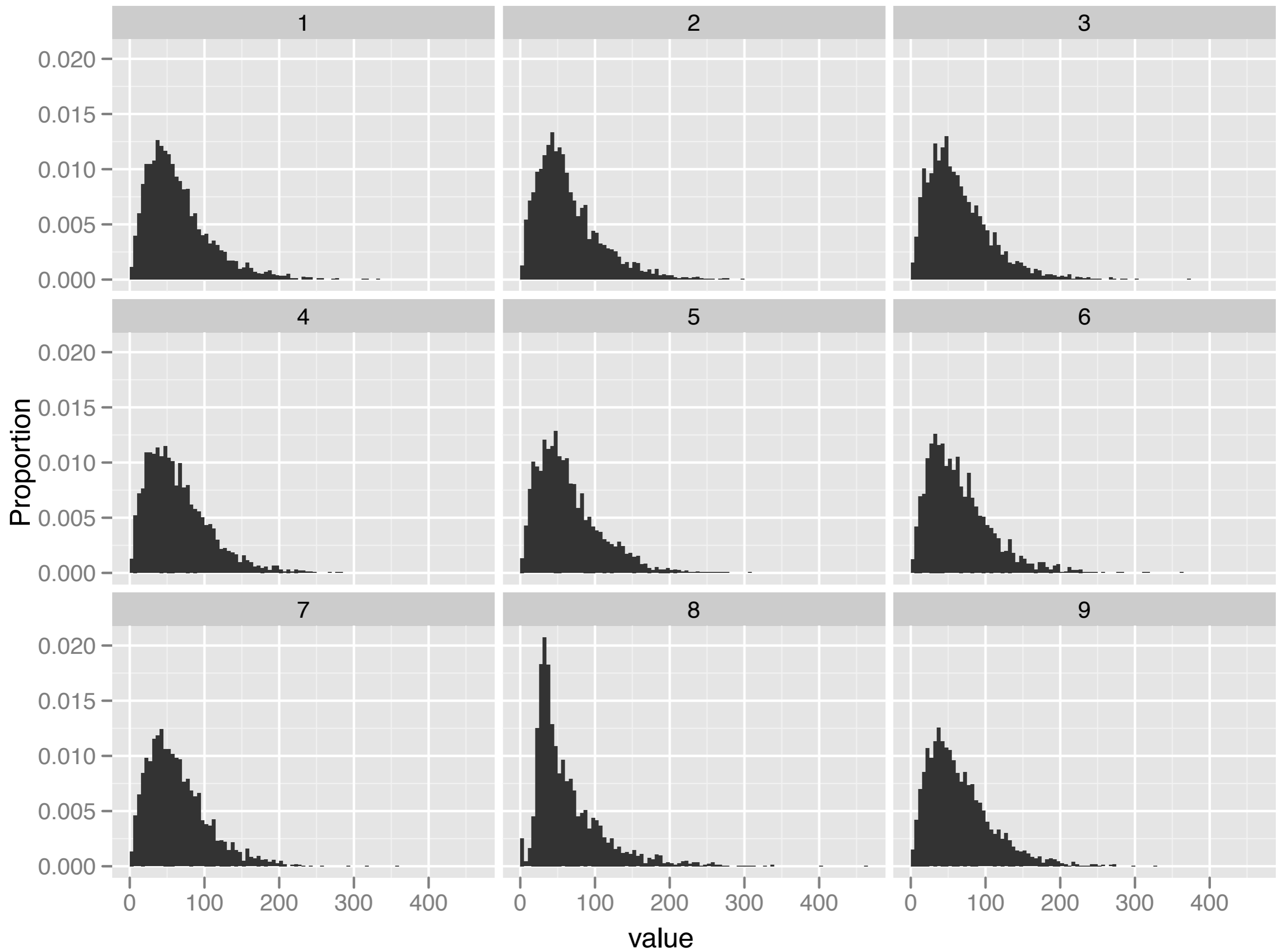
<p>Suspect</p>	<p>2.9 2.1 3.0 3.2 1.2 3.0 3.3 1.2 2.3 1.5</p>
<p>Crime</p>	<p>Not from Normal(1.5, 1)</p>
<p>Innocents</p>	<p>0.9 0.9 2.4 1.4 2.1 1.4 1.5 1.1 2.6 1.0 1.7 1.3 0.8 0.9 2.1 2.1 2.6 2.8 1.1 0.5 2.1 3.0 2.1 1.8 3.5 1.4 0.0 1.8 2.6 3.7 2.3 2.8 3.5 1.9 1.0 0.1 2.3 2.2 2.2 0.4</p>







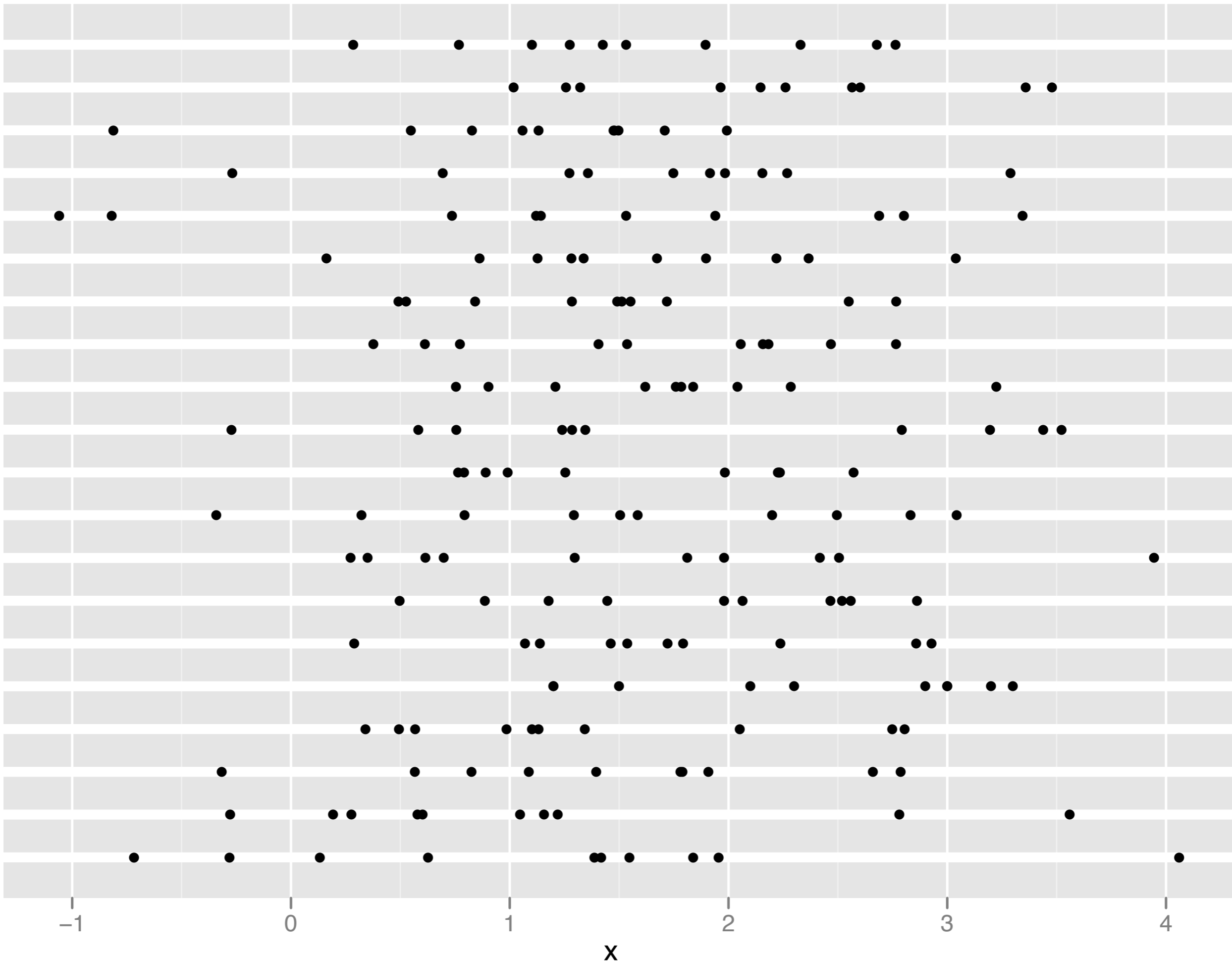




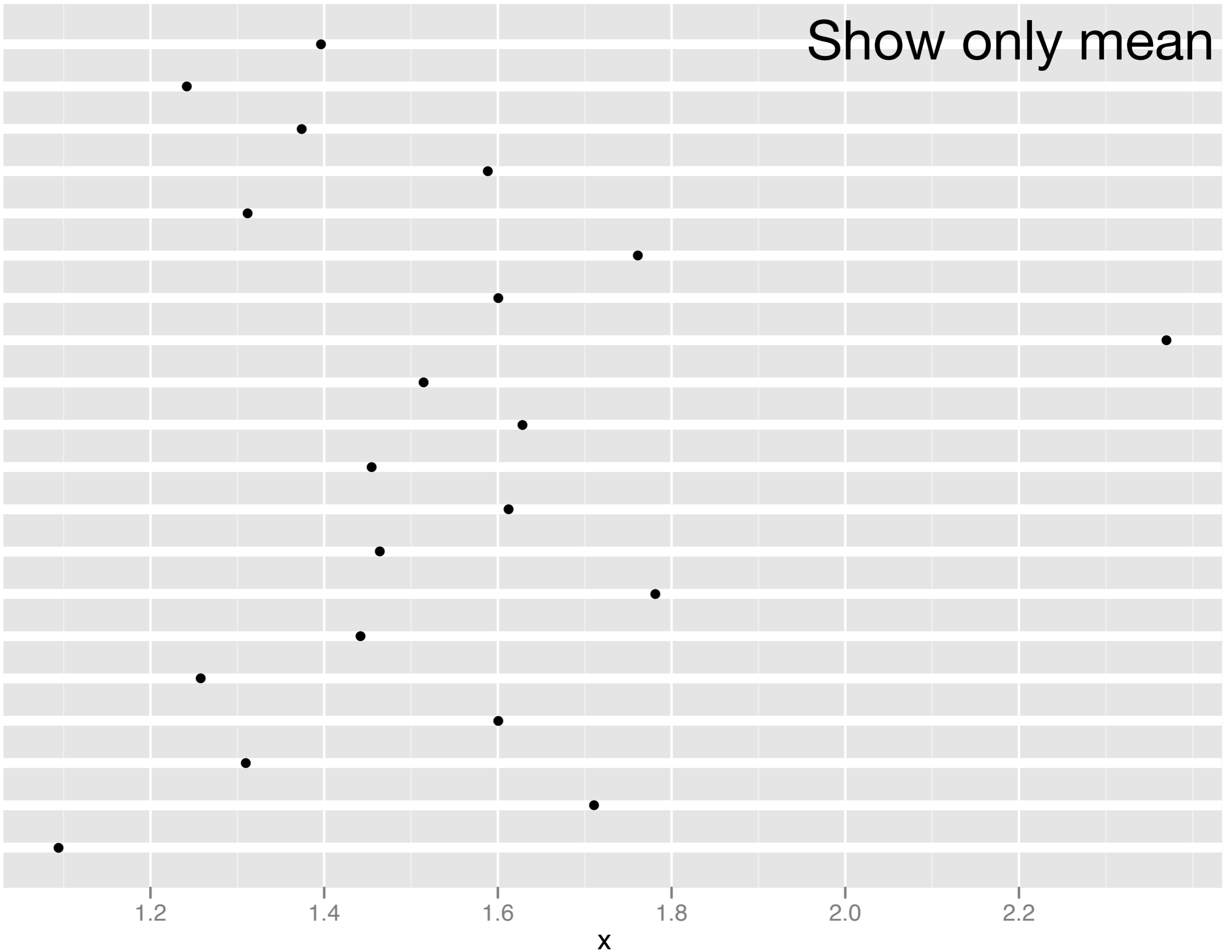
Test statistic

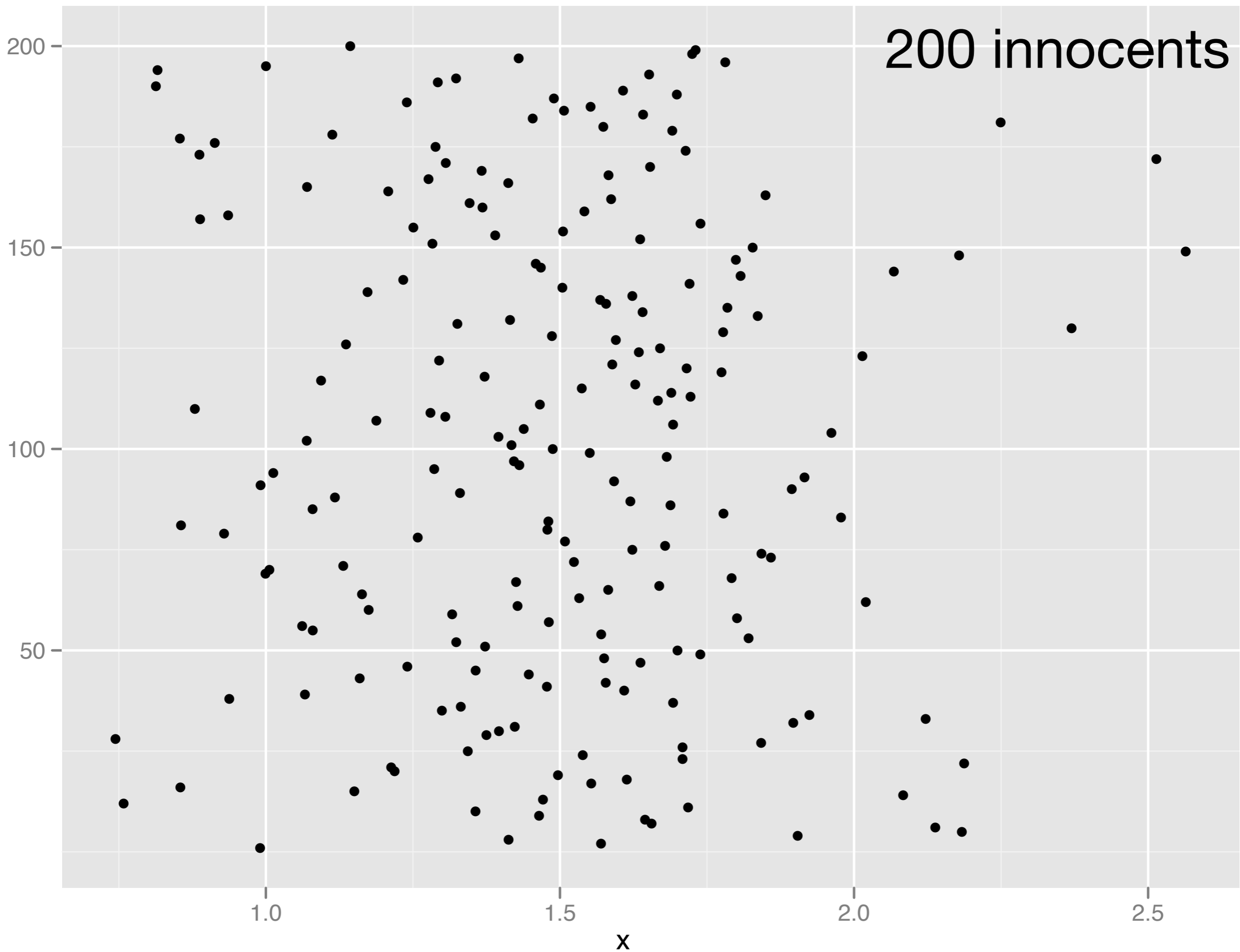
Need some way to put this on a more mathematical footing.

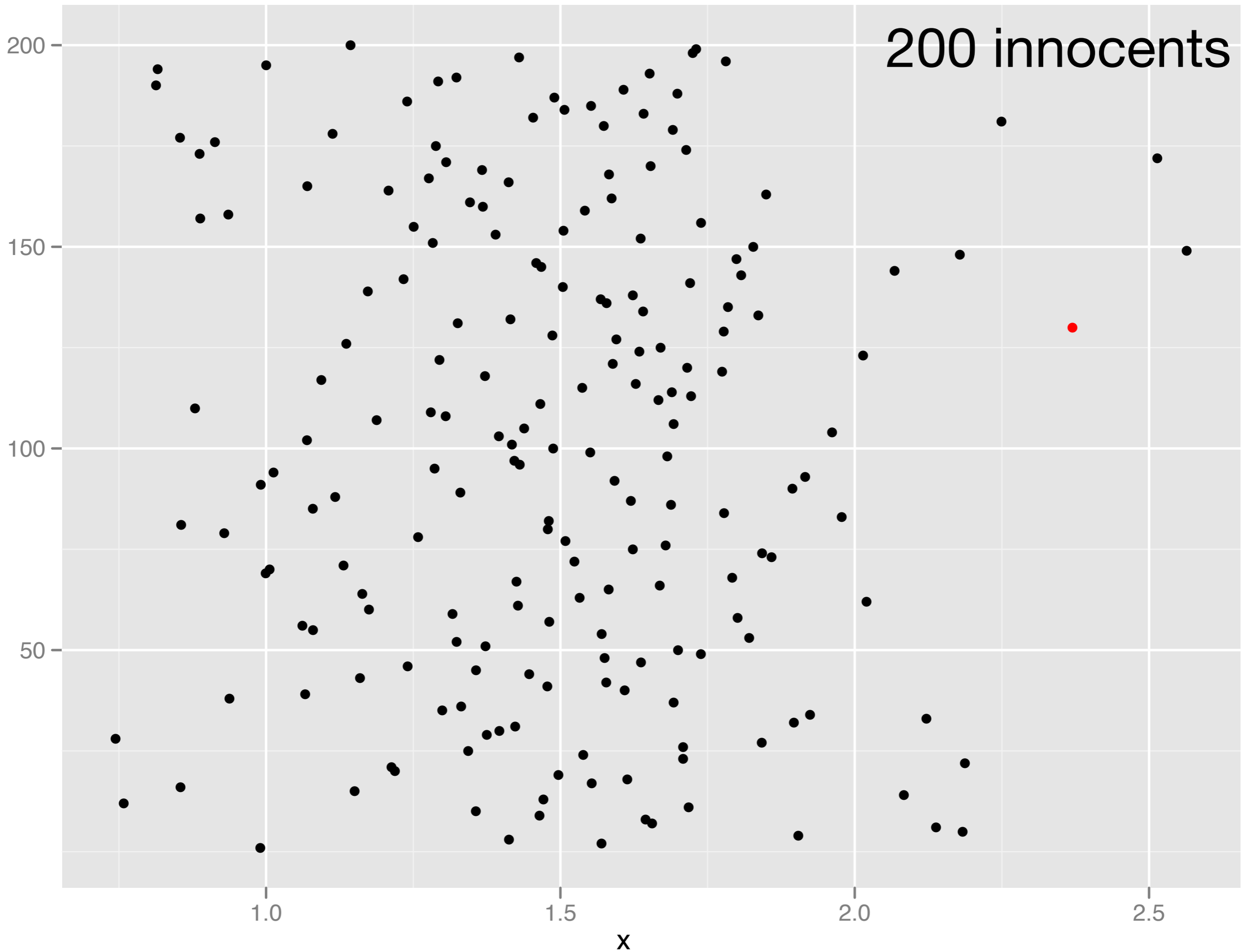
Need to summarise each innocent with a single number, and compare the suspect's summary to that number



Show only mean

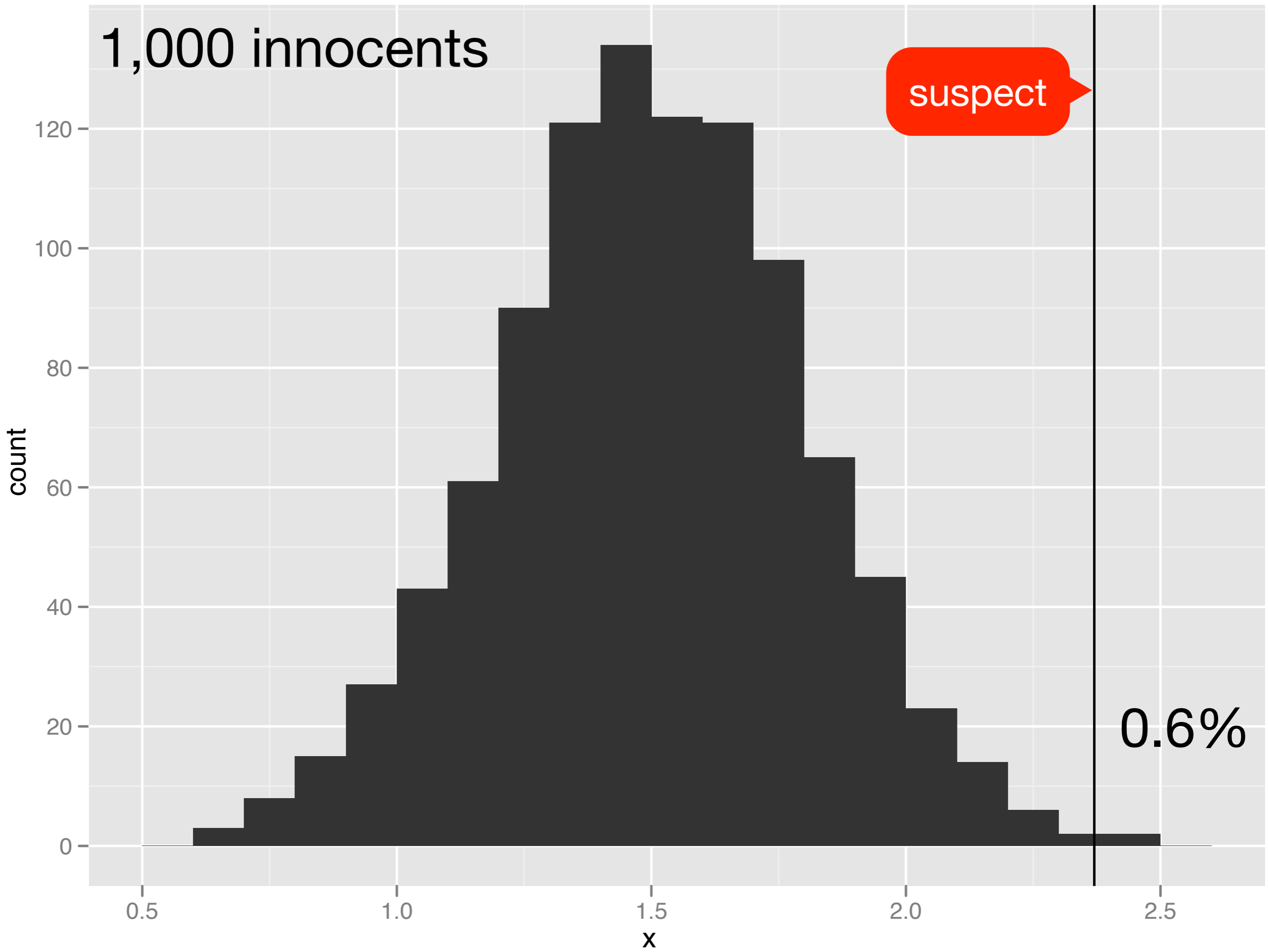






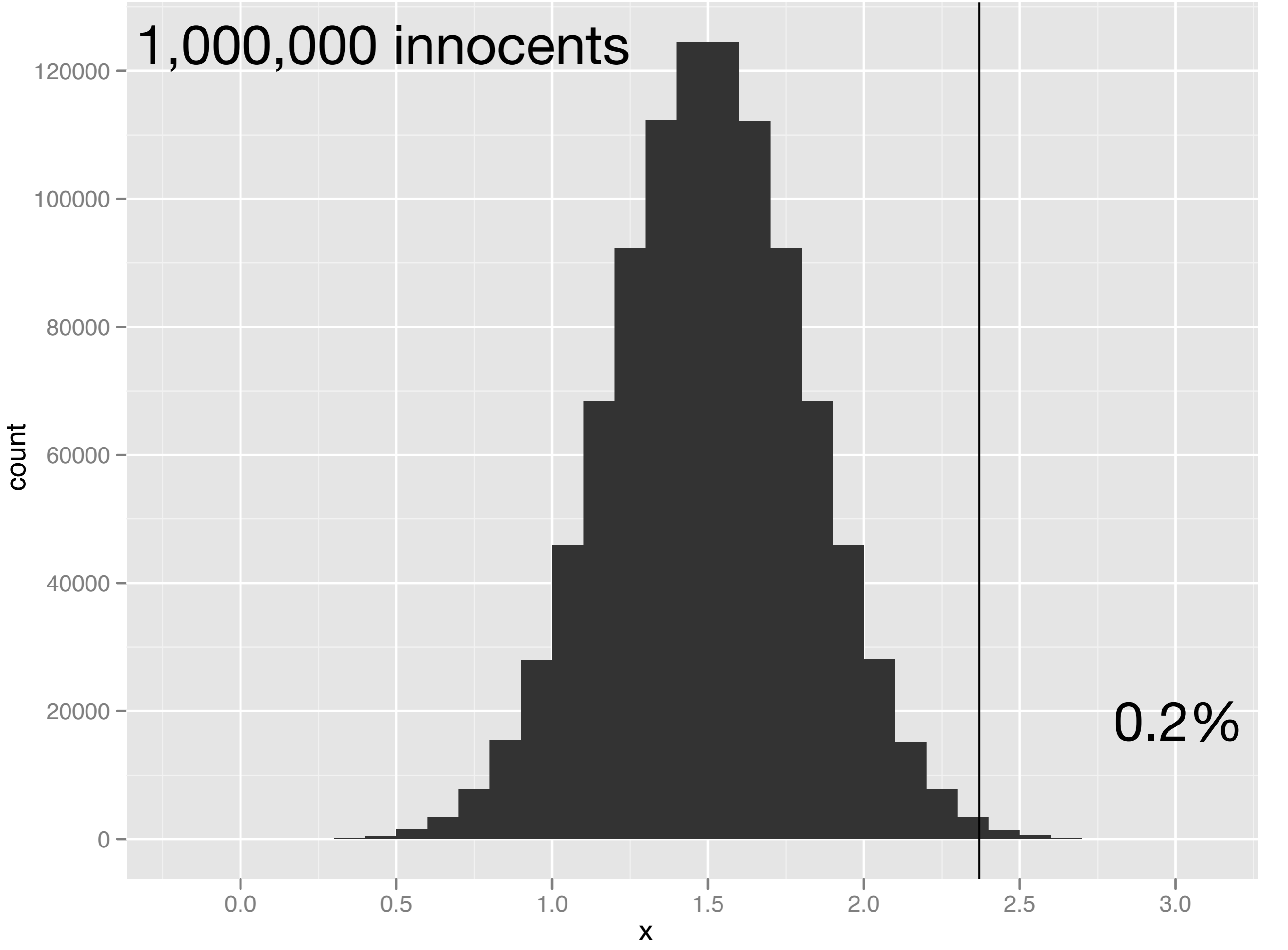
1,000 innocents

suspect



0.6%

1,000,000 innocents



0.2%

Your turn

What do we know about the distribution of innocents? What is the exact probability that an innocent would look as guilty as the suspect?

Would you convict or acquit?

(Mean of suspect is 2.37)

Your turn

The following values have generated from iid Normal(μ , 1):

2.9 2.1 3.0 3.2 1.2 3.0 3.3 1.2 2.3 1.5
(mean: 2.37)

Does $\mu = 1.5$?

Commonalities

- Innocent until proven guilty
- Never declare the suspect innocent:
only not guilty

Definitions

test statistic: numerical summary used to collapse sample into a single number

null distribution: distribution of test-statistic of innocents

p-value: probability that a true innocent would look as guilty as the suspect

Hypotheses

Every case has two sides:

The **null hypothesis** (the defence). This is the default, or the status quo, what you need to argue against.

The **alternative hypothesis** (the prosecution). This is the interesting case, but you need to prove it's true.

Rejecting the null

- To make a decision, we must decide on a threshold of the p-value: the **significance level**
- Never declare the suspect innocent: only not guilty
- **We fail to reject the null hypothesis, not accept the null hypothesis**

In the statistical justice system evidence is based on the similarity between the accused and known innocents.

The population of innocents, called the **null distribution**, is generated by the combination of null hypothesis and test statistic.

To determine the guilt of the accused we compare them to innocents. This is summarised with the **p-value**, the probability that a true innocent would look as (or more) guilty than the accused.

Case study

The lady tasting tea

A experiment by R. A. Fisher (famous early statistician, 1890-1962)

A lady at a tea party claims that she can tell the difference between putting the milk in first and second.

How can we be sure?

Experiment

8 cups. 4 milk first, 4 milk second.
Presented in random order.

What is the null hypothesis?
(What is the position of the defence?)

What test-statistic might we use?
(What can we measure to assess evidence?)

What is the null-distribution?
(What do innocents look like?)

Right	Wrong	#	%
4	0	1	1%
3	1	16	23%
2	2	36	51%
1	3	16	23%
0	4	1	1%
		70	100%

1. Write down null and alternative hypotheses (positions of defence and prosecution)
2. Figure out good test statistic (what numeric summary captures useful evidence)
3. Work out null distribution (distribution of innocents)
4. Calculate p-value by comparing actual value to null distribution (what proportion of true innocents look more guilty than the suspect)

Your turn

With a partner, write down the criminal equivalents to the following terms:

null hypothesis, alternative hypothesis, test statistic, null distribution, reject the null hypothesis, fail to reject the null hypothesis, accept the alternative hypothesis

Next time

- More practice identifying null and alternative hypotheses
- Development of specific tests from these general principles