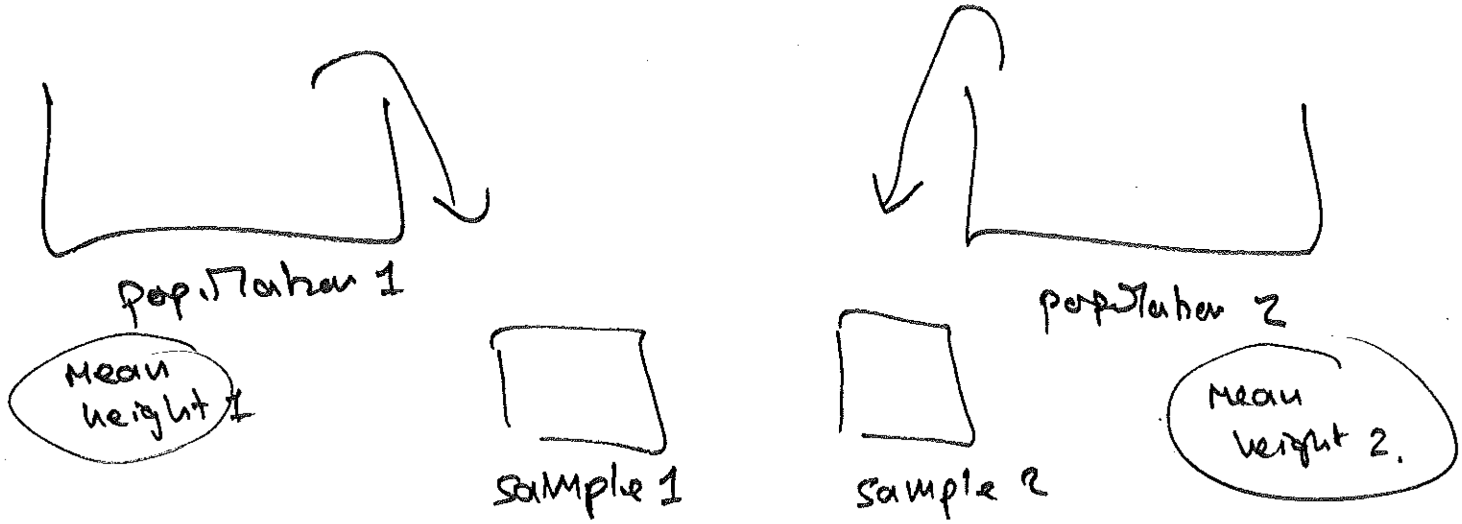


Based on the information in the two samples, can we conclude that the mean heights of the two populations are different?

Is the difference in the sample means consistent with the population means being the same, and chance error in the sample means due to sampling variability?

H_0 : mean height of population 1 = mean height of population 2.

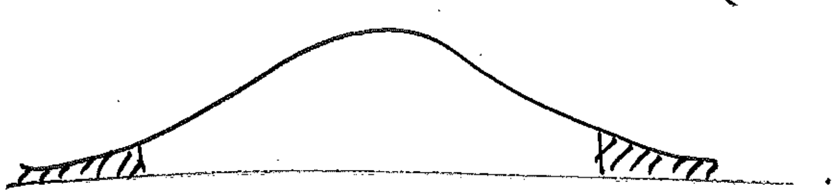
H_1 : mean heights of the 2 populations differ test statistic measures the size of the observed difference in the sample means in terms of the size of the difference we would expect (if H_0 were true, and the variability is just sampling variability)

$$Z = \frac{\text{observed difference} - \text{expected difference}}{SE_{\text{difference}}}.$$

$$SE_{\text{diff.}} = \sqrt{(1^{st}\ SE)^2 + (2^{nd}\ SE)^2}$$

SE_{mean} for sample 1.

p-value is the probability of observing a test statistic as large as the one we have, or larger, if H_0 is true. (= shaded area.)



if $p < 0.05$ we reject H_0, and say that the result is statistically significant.

if $p > 0.05$ we say that we do not have evidence to reject H_0.

What happens if there are more than 2 classes?

→ We have a hypothesis that specifies how frequently each of the classes ~~should~~ appears in the population.

How do we test the data in a sample to see if it is/is not compatible with the hypothesis?

eg Is a die fair?

Does the distribution of the number of times each face comes up correspond to that predicted by the null hypothesis?

$\boxed{1}\ \boxed{2}\ \boxed{3}\ \boxed{4}\ \boxed{5}\ \boxed{6}$

[Can't test each number separately as they are not independent]

value	observed frequency	expected frequency
1	4	10
2	6	10
3	17	10
4	16	10
5	8	10
6	9	10
Total	60	

Need a measure of how far the observed frequencies are from the expected frequencies.

$$\chi^2 = \text{sum of } \frac{\left(\text{observed frequency} - \text{expected frequency}\right)^2}{\text{expected frequency}}$$

(intuitively this has the right sort of properties)

$$\chi^2 = \frac{(4-10)^2}{10} + \frac{(6-10)^2}{10} + \frac{(17-10)^2}{10} + \frac{(16-10)^2}{10}$$

$$+ \frac{(8-10)^2}{10} + \frac{(9-10)^2}{10}$$

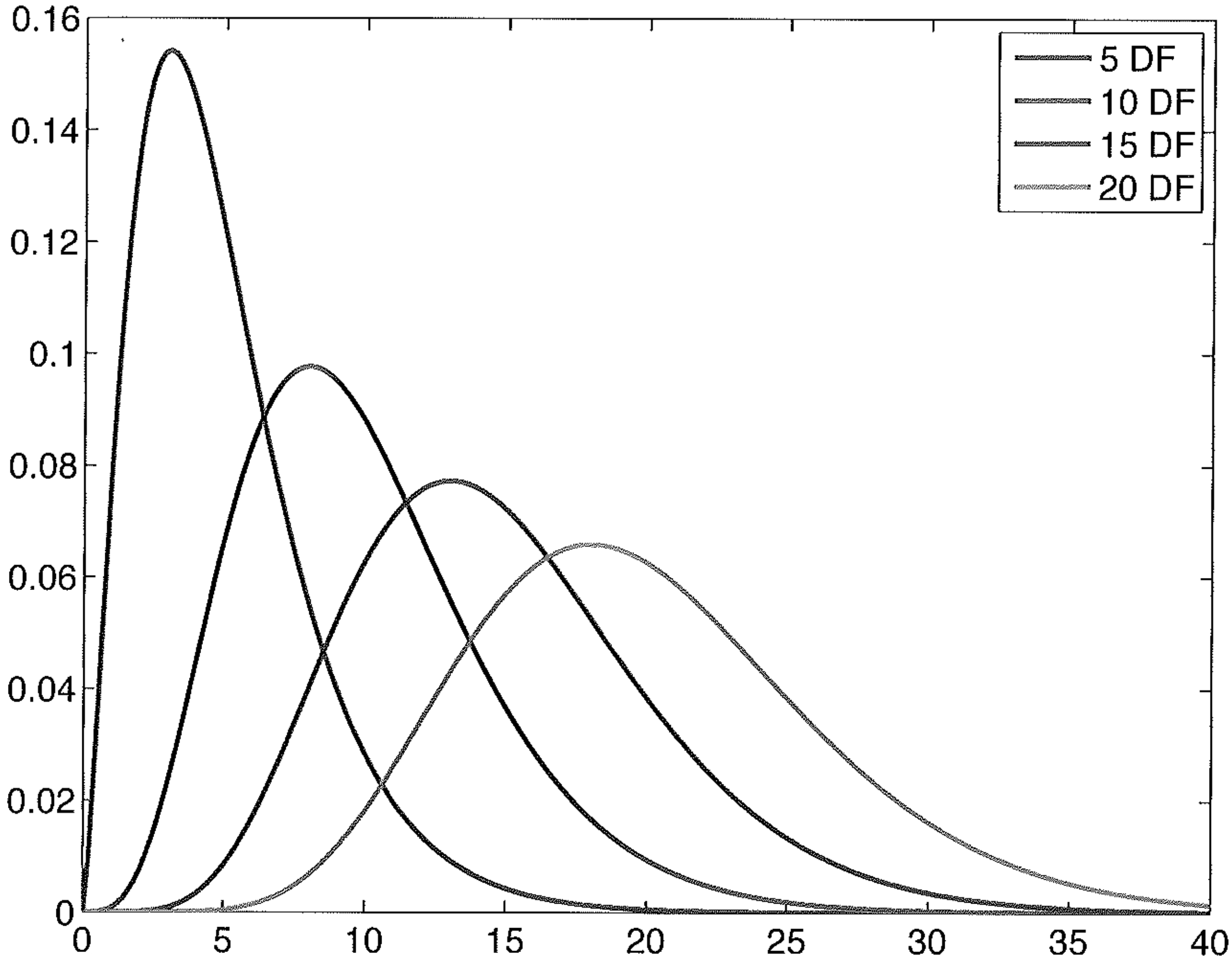
$$= \frac{142}{10} = \underline{\underline{14.2}}.$$

Now, under Ho (fair die), what is the chance of obtaining data that give a χ^2 value ≥ 14.2 ?

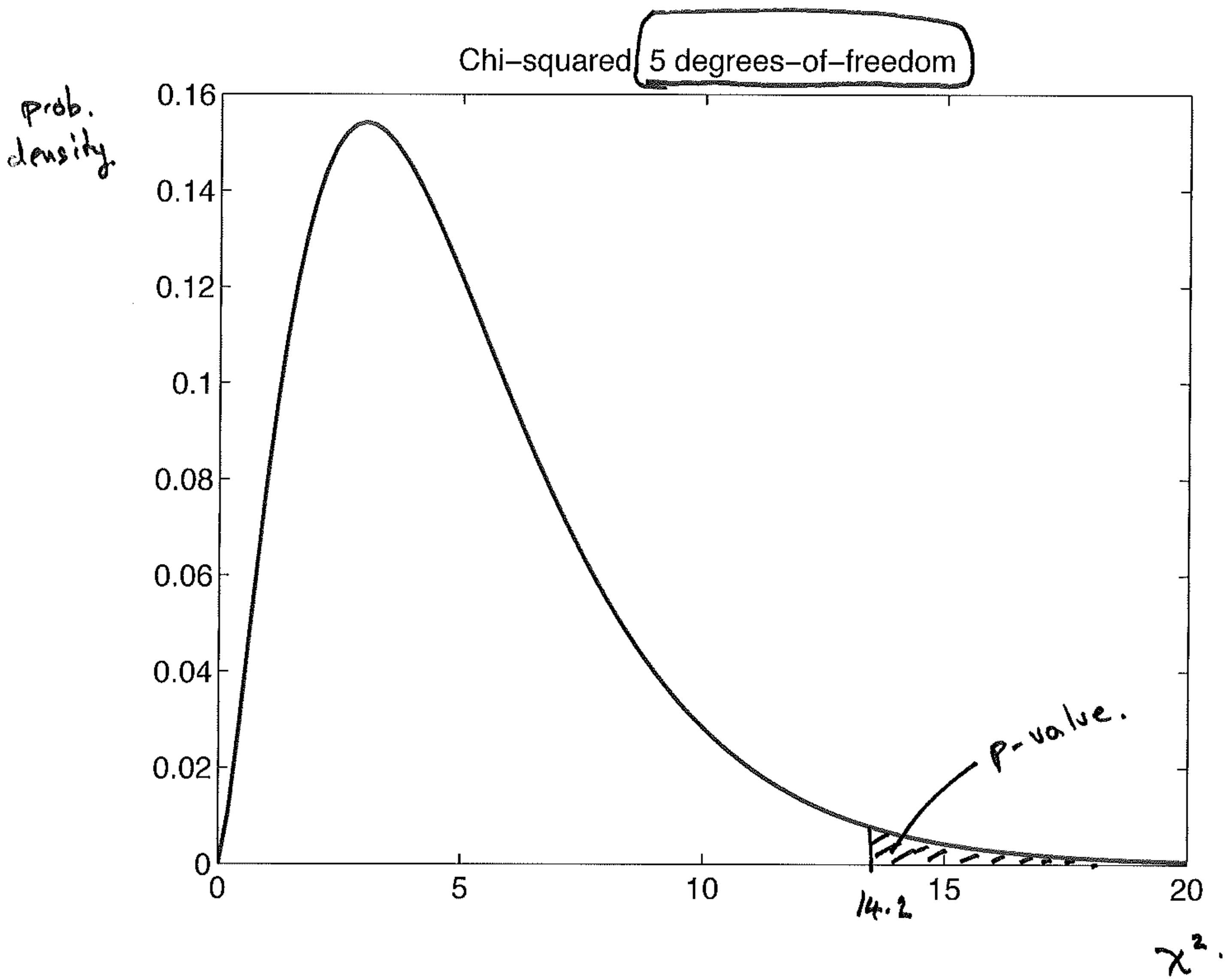
ie could we get a value like this just from sampling variability?

probability of getting different values for the χ^2 statistic depends on how many categories there are.

χ^2 curves are parameterized by the number of <u>degrees of freedom</u>.



5 DF
10 DF
15 DF
20 DF
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
20
25
30
35
40



Chi–squared 5 degrees–of–freedom
prob. density
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
20
14.2
p-value.
χ².

How to determine the number of DoF ?

Ho had specified the frequencies of each of the 6 possible outcomes.

However, knowing the frequencies of 5 of the outcomes + total # rolls, fixes the frequency of the 6^{th}.

In a <u>fully specified model</u> ← no parameters of the model that must be estimated from the data.

$$\text{DoF} = \begin{array}{c}\#\text{ terms in the}\\ \text{sum for } \chi^2\end{array} - 1$$

In this case $\#\text{DoF} = 6-1 = 5$

The p-value (chance that, if Ho is true, the χ^2 value computed from the data is ~~too~~ as large or larger than the one we have) is given by the area to the right of $\chi^2 = 14.2$ on the graph of χ^2 with 5 DoF.

Look these p-values up in a table...

from the table, the p-value is
between 5% and 1% (+ closer to 1%).

<u>So</u> : chance of getting $X^2 = 14.2$ or larger for
60 rolls of a fair die is $\simeq 1\%$.

<u>Reject</u> Ho that the data came
from a fair die.

<u>Note</u>.

- model is fully specified.
- expected frequency on
  each row of the table is large (≥ 5).

Day of week	observed # earthquakes	expected
M	16853	16916
T	16553	16916
W	16490	16916
Th	17399	16916
Fr	16348	16916
Sa	17019	16916
Sun	17753	16916
total	118,415	

Ho : earthquakes are equally likely to happen on each day of the week.

$$\chi^2 = \text{sum} \frac{\left(\text{observed freq} - \text{expected freq}\right)^2}{\text{expected freq.}} = 94.$$

6 Dof.

p-value << 1%

Reject Ho that earthquakes have no preference for particular days of the week.

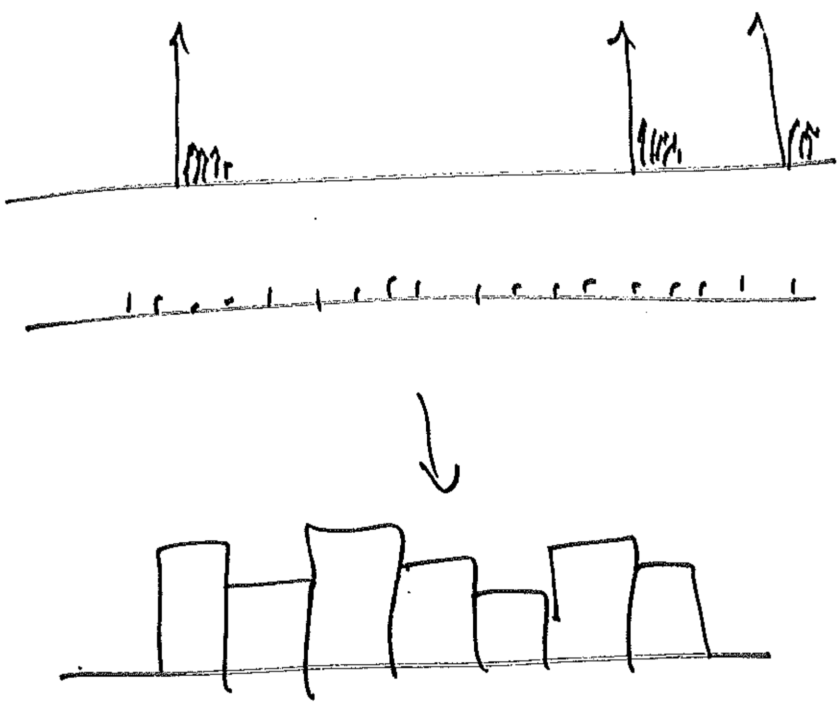
Does anyone believe this?

Total # earthquakes is large

large earthquakes is <u>small</u>

→ by sampling variability, there will be some days with more of them, and some with fewer.

After a large earthquake, lots of aftershocks.



large earthquakes + aftershocks

lots of small earthquakes

non-uniform distribution.

What's the point here?

Rejecting H_0 is the start of a process of thinking a little more carefully about what the process actually is that generated the data.

Testing Independence

	M	F
Rt handed		
left handed		
amb.		

are ~~is~~ handedness and sex independent?

Is having facial hair independent of the department of the hospital that you work in?

H_0 : independence.

What are the expected values predicted by H_0 ?

Box model.

$\boxed{\text{RH Man}}$ $\boxed{\text{RH Woman}}$

$\boxed{\text{LH Man}}$ $\boxed{\text{LH Woman}}$

$\boxed{\text{Amb. Man}}$ $\boxed{\text{Amb. Woman}}$

H_0: handedness and sex are independent.

% of RH amongst men equals % RH amongst women etc.

H_1: distribution of handedness differs between men and women.

	M	F
R	26	42.
L	4	7
A	1	3.

	M	F	
Rt handed	26	42	68
not R. handed	5	10	15
	31	52	83

<u>observed</u>

<u>independence</u>: % of men + women on each <u>row</u> should be the same.

$\frac{68}{83}$ is the fraction of all people who are Rt handed.

Since we have 31 men, we would expect $31 \times \frac{68}{83}$ Rt handed men.

we have 52 women, we would expect $52 \times \frac{68}{83}$ Rt handed women.

In general:

$$\text{expected value in a cell} = \frac{\text{row total} \times \text{column total}}{\text{table total.}}$$

expected values.

	M	F		
Rt	$\frac{68\times31}{83}$	$\frac{68\times52}{83}$	25.4	42.6
not Rt	$\frac{15\times31}{83}$	$\frac{15\times52}{83}$	5.6	9.4.

$$\chi^2 = \text{sum} \frac{\left(\text{observed freq} - \text{expected freq}\right)^2}{\text{expected freq.}}$$

$$= \frac{(26-25.4)^2}{25.4} + \frac{(42-42.6)^2}{42.6} + \frac{(5-5.6)^2}{5.6} + \frac{(10-9.4)^2}{9.4}$$

$$= \underline{\underline{0.125}}.$$

How many degrees of freedom?

Don't know the true % of each ticket -
estimating them from the data.

- lose 1 Dof for each parameter we estimate.

Table with m rows and n colums
has $(m-1) \times (n-1)$ dof.

In this example $m = 2$, $n = 2$ $\Rightarrow (2-1) \times (2-1) = 1$ dof.

From the table, the p-value is a little bit less than 90%.

Do not reject the null hypothesis that sex and handedness are independent.