Lecture 4 - Bayesian correlation

Let's recall the definition of the (Pearson) correlation coefficient.

Consider two sets of scores - to what degree are they associated?

↳ i.e., how do they co-vary?

X	Y	$X-\bar{X}$	$Y-\bar{Y}$	$(X-\bar{X})(Y-\bar{Y})$
6	6	2	2	4
2	2	-2	-2	4
5	6	1	2	2
3	4	-1	0	0
4	2	0	-2	0
$\bar{X}=4$	$\bar{Y}=4$			
$\sigma_X=1.41$	$\sigma_Y=1.79$			

Define: covariance

$$\sigma_{XY} = \frac{1}{N}\sum(X_i-\bar{X})(Y_i-\bar{Y})$$

$$\sum = 10 \longrightarrow \sigma_{XY} = \frac{1}{5}\times 10 = 2$$

From covariance, we can define the Pearson correlation:

$$\boxed{\rho = \frac{\sigma_{XY}}{\sigma_X\cdot\sigma_Y}}$$

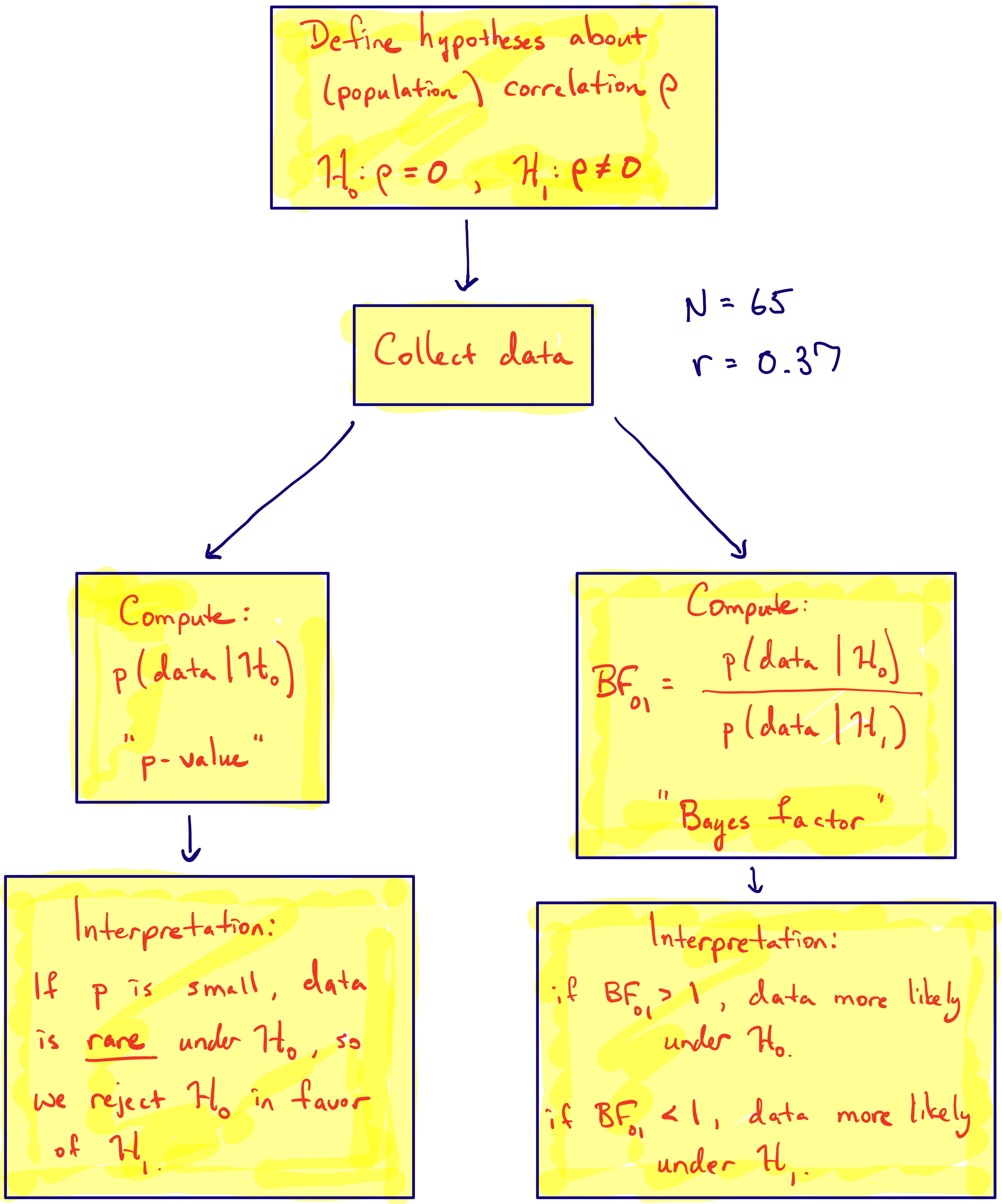
So for our example,

$$\rho = \frac{2}{(1.41)(1.79)} = 0.79$$

Facts:

- $* \quad -1 \le \rho \le 1$
- $*$ as $\rho \to \pm 1$, degree of association increases

Suppose we are interested in the relationship between math anxiety and performance on a standardized assessment.



Define hypotheses about (population) correlation ρ

$\mathcal{H}_0 : \rho = 0$, $\mathcal{H}_1 : \rho \neq 0$

Collect data

$N = 65$

$r = 0.37$

Compute:

$p(\text{data} \mid \mathcal{H}_0)$

"p-value"

Interpretation:

If p is small, data is rare under $\mathcal{H}_0$, so we reject $\mathcal{H}_0$ in favor of $\mathcal{H}_1$.

Compute:

$$BF_{01} = \frac{p(\text{data} \mid \mathcal{H}_0)}{p(\text{data} \mid \mathcal{H}_1)}$$

"Bayes factor"

Interpretation:

if $BF_{01} > 1$, data more likely under $\mathcal{H}_0$.

if $BF_{01} < 1$, data more likely under $\mathcal{H}_1$.

p-value = $p(\text{data} \mid H_0)$

1) only considers fit of H_0 as a potential model for data

2) ignores fit of H_1

Thus, "support" for H_1 is only indirect

Bayes factor = $\dfrac{p(\text{data} \mid H_0)}{p(\text{data} \mid H_1)}$

1) considers relative adequacy of both models as predictors of data.

2) can directly index support for either H_0 or H_1.

Ex: $BF_{01} = 8$ → "The observed data are 8 times more likely under H_0 than H_1."

Jeffreys (1961):

BF	Evidence*
1 - 3	anecdotal
3 - 10	moderate
10 - 30	strong
30 - 100	very strong
> 100	extreme

$BF_{10} = \dfrac{1}{8}$

* these are only guidelines!

How does Bayes work?

for single model $\mathcal{H}$:

$$\boxed{p(\mathcal{H} \mid \text{data}) = p(\mathcal{H}) \times \frac{p(\text{data} \mid \mathcal{H})}{p(\text{data})}}$$

$$\underset{\text{belief in } \mathcal{H}}{\text{posterior}} = \underset{\text{belief in } \mathcal{H}}{\text{prior}} \times \text{updating factor}$$

for <u>two</u> models:

$$\boxed{\frac{p(\mathcal{H}_0 \mid \text{data})}{p(\mathcal{H}_1 \mid \text{data})} = \frac{p(\mathcal{H}_0)}{p(\mathcal{H}_1)} \times \frac{p(\text{data} \mid \mathcal{H}_0)}{p(\text{data} \mid \mathcal{H}_1)}}$$

$$\underset{\text{odds}}{\text{posterior}} = \underset{\text{odds}}{\text{prior}} \times \text{Bayes factor}$$

What do we mean by prior?

Two types of "priors":

1) priors on models

2) priors on parameters within a given model

① Priors on models — before observing data, what is relative likelihood of competing models?

- common default: $p(\mathcal{H}_0) = p(\mathcal{H}_1) = 1/2$

  ↳ i.e., "1-1 prior odds"

- these prior model probabilities must add to 1

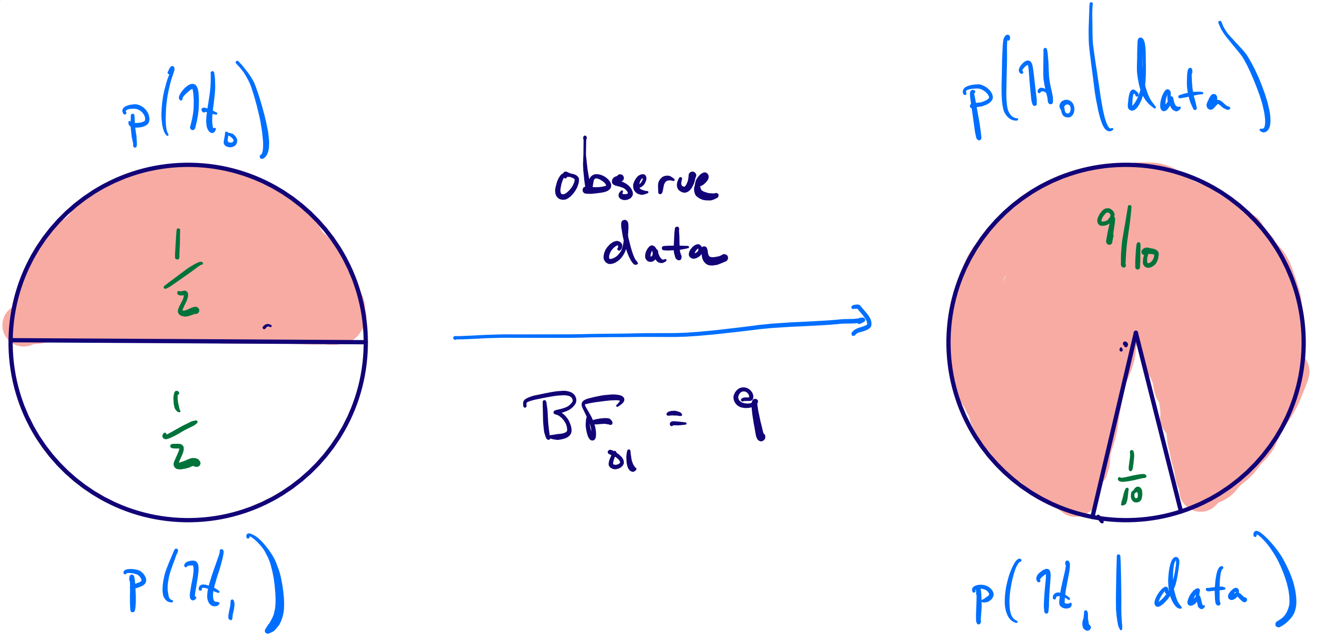
  ↳ $p(\mathcal{H}_0) + p(\mathcal{H}_1) = \frac{1}{2} + \frac{1}{2} = \underline{1}$

- prior model probabilities are updated after observing data:

Posterior odds = prior odds × BF

$$\boxed{p(\mathcal{H}_0 \mid \text{data}) = \frac{\text{posterior odds}}{1 + \text{posterior odds}}}$$

Example:



Prior odds = 1:1

Posterior odds = 9:1

$$p(\mathcal{H}_0 \mid \text{data}) = \frac{\text{posterior odds}}{1 + \text{posterior odds}}$$

$$= \frac{9}{9+1} = \frac{9}{10} = 0.90$$

② priors on parameters within a given model.

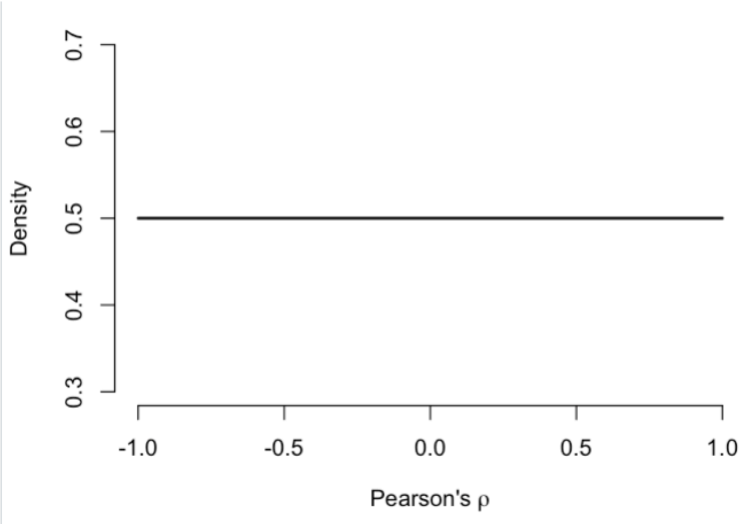
- model definitions : $\mathcal{H}_0 : \rho = 0$

  $\mathcal{H}_1 : \rho \neq 0$ ← what exactly do we mean here?

- we quantify our uncertainty about the correlation ρ under $\mathcal{H}_1$ by placing a distribution on ρ

- suppose we have no idea what to expect. Here, we might believe any value of ρ is equally likely to occur.

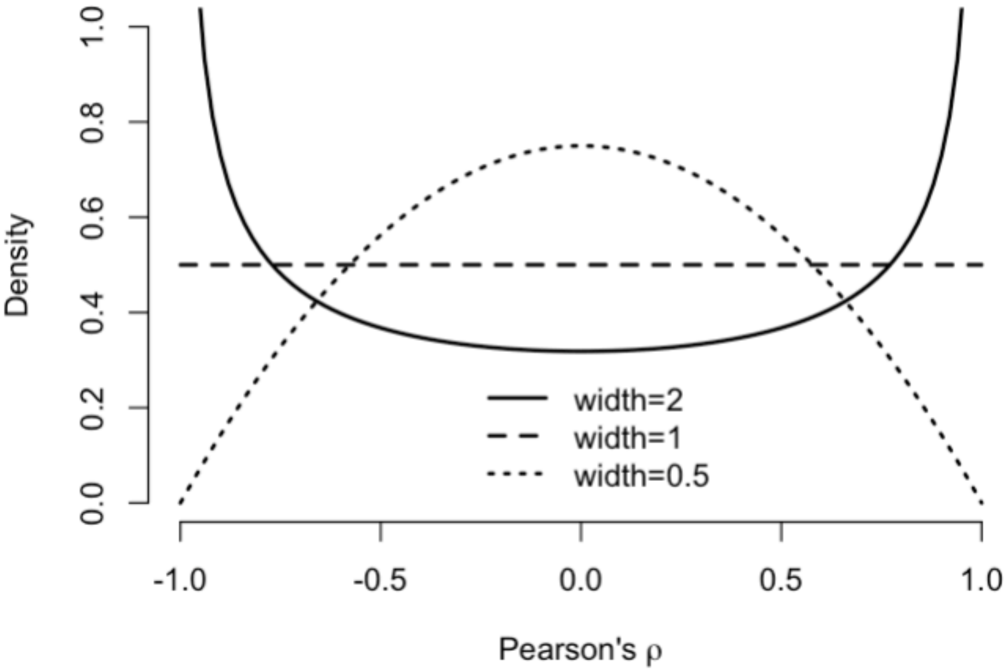
  ↳ we say ρ is uniformly distributed on $(-1, 1)$



* Note - this is the "default" used in JASP.

Other priors:

↳ the uniform prior is a special case of a useful class of priors called a "stretched Beta prior"



Shape of the prior is controlled by the width parameter

* width 1 → uniform prior (JASP default)

* width 0.5 → more prior mass at 0

↳ useful if you believe the correlation is small

* width 2 → more prior mass at ± 1

↳ useful if you believe correlation is strong (ex: testing reliability)

Let's continue our working example. Suppose we tested $N = 65$ participants and observed a correlation of $r = 0.37$.

- use JASP "Summary Statistics" module

Elements to report:

1. report results of hypothesis test
   - define H_0, H_1 and specify prior under H_1.

     "Under the null hypothesis we expect a correlation of 0 between maths anxiety and performance. Thus, we define $H_0: \rho = 0$. The alternative hypothesis is two-sided, $H_1: \rho \neq 0$, and we assigned a uniform prior probability to all values of ρ between -1 and $+1$."

   - report and interpret Bayes factor

     "We found a Bayes factor of $BF_{10} = 13.93$, which means that the observed data are approximately 14 times more likely under H_1 than H_0. This result indicates strong evidence in favor of H_1"

- calculate and report posterior model probability for preferred model.

  - from earlier,

$$p(\mathcal{H}_1 \mid \text{data}) = \frac{\text{posterior odds}}{1 + \text{posterior odds}}$$

$$= \frac{13.93}{1 + 13.93} = 0.93.$$

- "Assuming prior odds of 1-1 for $\mathcal{H}_1$ and $\mathcal{H}_0$, our observed data updated these odds to 13.93-to-1 in favor of $\mathcal{H}_1$. This is equivalent to a posterior model probability of $p(\mathcal{H}_1 \mid \text{data}) = 0.93$."

2. report results of parameter estimation
  - only if $\mathcal{H}_1$ is the preferred model!
  - specify parameter of interest and remind reader of prior under $\mathcal{H}_1$
    - "Of interest is the posterior distribution for ρ, the population-level correlation between maths anxiety and performance. Under $\mathcal{H}_1$, ρ was assigned a uniform prior over the interval from -1 to $+1$."

- report the 95% credible interval.
    - "The posterior distribution for ρ had a median of 0.356, with a central 95% credible interval that ranges from 0.134 to 0.554."