Some Bayesian Inference

Bayesian Data Analysis

Steve Buyske

Some Inference Part 1: the Probability of Direction

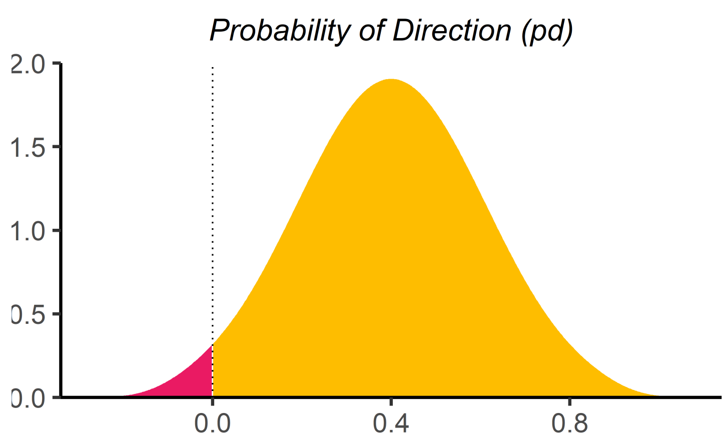
- You may have noticed that the `describe_posterior()` function has a column labeled `pd`.

```
Parameter   | Median |           90% CI |      pd |        90% ROPE | % in ROPE |  Rhat |      ESS
-------------------------------------------------------------------------------------------------
(Intercept) | 87.986 | [83.621, 92.202] | 100.00% | [-0.717, 0.717] |         0 | 1.001 | 3894.816
gini        | -0.388 | [-0.492, -0.275] | 100.00% | [-0.717, 0.717] |       100 | 1.001 | 3981.026
```

- `pd` stands for *Probability of Direction*, also known as *Maximum Probability of Effect*.

Probability of Direction

- It lies between 50% and 100%
- It is the probability (based on the posterior) that a parameter is strictly positive or negative, whichever is the most probable.
- You could calculate it by first looking at the sign of the median of the posterior, and then calculating the probability that the parameter is on the same side of zero.
- The area in yellow is the probability of direction in this illustration.



Probability of Direction cont.

- Notice that the probability of direction doesn't say anything about the magnitude of a parameter.
- It is more a statement about the probability that it lies to one side of zero.
    - That is, that is *has* a direction.

- The probability of direction is very similar to what people incorrectly think the p-value is telling them, although in this case it would be what $1 - p/2$ is telling them, where p is a two-sided p-value.
- A two-sided p-value of .1, .05, .01 and .001 would roughly correspond to a pd of 95%, 97.5%, 99.5% and 99.95%, respectively.

- In R, if all you want is the probability of direction, you can get it with `p_direction()`, where you put the name of the object (i.e, `gini_stan`) inside the parentheses.

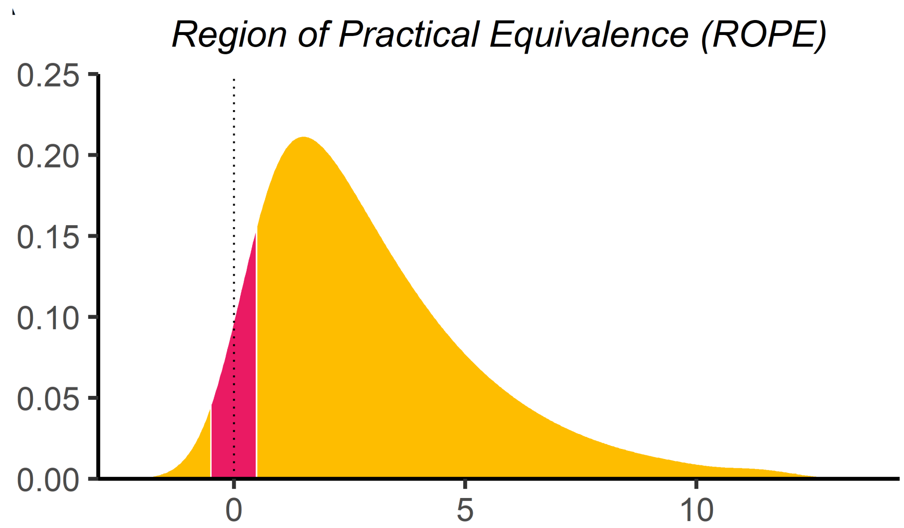
```
p_direction(gini_stan)
```

```
## # Probability of Direction (pd)
##
## Parameter   |      pd
## --------------------
## (Intercept) | 100.00%
## gini        | 100.00%
```

- This indicates virtual certainty in this example that the intercept term and the regression coefficient for `gini` (that is α and β) are away from zero.

Some Inference Part 2: ROPE

- Of course, we may not be interested if most of the probability of θ is positive, say, but very, very small.
- The idea of a region of practical equivalence, or ROPE, is that you would define a region around zero which is equivalent in a practical sense to zero.
    - One might then calculate the probability that θ is outside that region.
        - That probability would be the area in yellow below.



- Instead, however a popular approach is the *HDI + ROPE* decision rule:
  - If the HDI (the 90% HDI, say) is completely outside the ROPE, then you *reject* the idea that the parameter is zero.
    - "… is equivalent in a practical sense to zero" would be more precise.
    - The use of "reject" is meant to be analogous to hypothesis testing in the frequentist framework.
  - If the HDI is completely inside the ROPE, then you *accept* the idea that the parameter is zero.
  - Otherwise, you are *undecided*.

- You can instead quantify this information by calculation the percentage of the HDI that is inside the ROPE.

```
Parameter   | Median |          90% CI |      pd |       90% ROPE | % in ROPE |  Rhat |      ESS
-------------------------------------------------------------------------------------------------
(Intercept) | 87.986 | [83.621, 92.202] | 100.00% | [-0.717, 0.717] |         0 | 1.001 | 3894.816
gini        | -0.388 | [-0.492, -0.275] | 100.00% | [-0.717, 0.717] |       100 | 1.001 | 3981.026
```

- Incidentally, the use of ROPE is much more common in fields near to psychology than in fields more distant from it—the developer of the concept is a psychologist.

How big a ROPE?

- How big should the region of practical equivalence be?
  - If you have knowledge of the field, you can use that.
    - For example, I've read that a change in blood pressure of 3 mmHg is considered clinically meaningful, so you might have a ROPE of $(-3, 3)$.
  - Otherwise, a common suggestion is to use $(-0.1 \times \sigma_y, 0.1 \times \sigma_y)$, where σ_y is the standard deviation of the outcome variable.
    - Seems to make more sense if you are interested in means than for a regression setting though.

- The function `describe_posterior()` will use a default for the ROPE of $(-0.1 \times \sigma_y, 0.1 \times \sigma_y)$ and show the percentage overlap of the ROPE and the HDI.
    - The default size of the HDI is 89%; to change it to 90% use the `rope_ci` argument.

```
describe_posterior(gini_stan, ci = 0.9, rope_ci = 0.9)
```

```
## # Description of Posterior Distributions
##
## Parameter   | Median |           90% CI |      pd |        90% ROPE | % in ROPE |  Rhat |      ESS
## --------------------------------------------------------------------------------------------------
## (Intercept) | 87.986 | [83.621, 92.202] | 100.00% | [-0.717, 0.717] |         0 | 1.001 | 3894.816
## gini        | -0.388 | [-0.492, -0.275] | 100.00% | [-0.717, 0.717] |       100 | 1.001 | 3981.026
```

- You can even get a decision made for you by using the `equivalence_test()` function.

```r
equivalence_test(gini_stan, ci = .9) # Notice that for this function the argument is ci, not rope_ci
```

```
## # Test for Practical Equivalence
##
##   ROPE: [-0.72 0.72]
##
## Parameter   |       H0 | inside ROPE |       90% HDI
## ----------------------------------------------------
## (Intercept) | Rejected |      0.00 % | [83.62 92.20]
## gini        | Accepted |    100.00 % | [-0.49 -0.27]
```

My own opinion

- I find the probability of direction and the ROPE useful for quick answers
  - What is the probability that the parameter lies on one side of zero?
    - The value of “pd”.
  - What is the probability that the parameter is practically different from zero?
    - The value of “% in ROPE”.
- However, the pd and the ROPE also seem like they were designed to make frequentists comfortable.
  - Why not just ask questions directly?
    - What is $\mathrm{Prob}(\theta > 0)$?
    - What is $\mathrm{Prob}(\theta < -2)$?

Some Inference Part 3: Direct calculation of probabilities from the posterior

- Since we have a large number of draws from the posterior, we can use those to answer questions about probabilities about a parameter.
- A question about the probability of a parameter satisfying some condition is approximately the same as a question about the proportion of draws of that parameter satisfying the same condition.

- For example, if we write θ for the parameter and $\{\theta_i\}$ for the set of draws from the posterior of θ, then
  - $\text{Prob}(\theta > 0)$ is approximated by the proportion of $\{\theta_i\} > 0$.
  - The more draws, the more accurate the approximation
    - This is particularly important if you are interested in small probabilities.

In R

- This is going to take a few steps to explain.
- After we fit a model, the object (like `exam_hier`) has all sort of pieces, but we want only the draws from the posterior)
- One way to do that is with the function `as.data.frame()`, which will give a dataframe of the draws.
    - Each column corresponds to all the draws from the posterior for that parameter.

```
exam_hier %>%
  as.data.frame()
```

```
##       (Intercept)  standLRT b[(Intercept) school:1] b[standLRT school:1]
## 1   -1.677100e-02 0.5357654              0.55441258         0.0985912778
## 2   -2.715209e-02 0.5493918              0.31892804         0.1348899429
## 3   -6.141010e-02 0.5564134              0.41806451         0.0764489116
## 4   -1.340228e-02 0.5655751              0.36562274         0.0997278122
## 5    1.308479e-02 0.5649491              0.40565355         0.1080196490
## 6    1.304463e-02 0.5641389              0.31249081         0.0825394729
## 7   -3.903748e-03 0.5534762              0.40185737         0.1846493038
## 8   -3.290939e-02 0.5759107              0.39646130         0.1086346649
```

If we want to do work with a particular variable we can use a special pipe `%$%`

```
exam_hier %>%
  as.data.frame() %$%
  mean(standLRT)
```

```
## [1] 0.5564451
```

- Remember that we are looking at draws from the posterior, so the variable `standLRT` here is does not represent the variable `standLRT` from the original data, but the regression coefficient β in front of `standLRT`.
  - It can be a little confusing.
- That code above gives us the mean of the draws.
  - Even though it is a number between 0 and 1, it’s not related to probability.

Finally, a probability

- To answer a probability question, we want the proportion of the draws satisfying some condition.
- In R, if you have a logical statement, R will turn that into `FALSE` or `TRUE`.
- If you use the logical statement in a mathematical expression, then R will turn it into a `0` or `1`.
- This means the proportion satisfying a condition is the same as the mean of that condition:

```
exam_hier %>%
  as.data.frame() %$%
  mean(standLRT > 0.6)
```

```
## [1] 0.01525
```

- Thus we find the probability that the regression coefficient for standLRT is greater than 0.6 is 0.01525.

Two last notes

- For variables that are not legal R names, like `(Intercept)` (since it starts with a parenthesis), you can enclose it in back-ticks:

```
exam_hier %>%
  as.data.frame() %$%
  mean(`(Intercept)` < 0)
```

```
## [1] 0.60425
```

- If you want to combine logical conditions, put parentheses around them and combine them with `&` for the logical *and*, `|` for the logical *or*.

```
exam_hier %>%
  as.data.frame() %$%
  mean( (`(Intercept)` < 0) & (standLRT > 0.6))
```

```
## [1] 0.00475
```