Modernizing *Into to Statistics* with *Modern Dive*

Lydia Gibson & Zhaoshan Duan

January 20, 2022

In this blog, we advocate for a modernized approach to introduce statistical inference in entry-level statistic courses. We plan to compare two ways of conducting hypothesis testing: base R vs. tidy with `infer` using some of the most common topics in *Intro to Statistics* courses.

Hypothesis testing forms the foundation of inferential statistics and is often introduced early in most *Intro to Statistics* classes. As students, we were overwhelmed by the various kinds of statistical tests, distributions, p-value, and their interpretations. It only gets better once students acquire the realization and intuition of the universal pattern that statistical tests follow through advanced courses.

Moreover, simulation-based inferences are often more visual and intuitive than analytical inferences. While many educators include those numerical approaches in their teaching, it often feels segmented, especially in base R code.

We believe that the `infer` framework embodies the intuition of "there's only one test", and presents hypothesis testing more coherently and consistently than base R for beginners in Statistics and/or R. We hope this blog post could provide students' perspectives on statistical education and shed light on the discussion of moderndizing "Intro to Statistics" courses.

Last Updated: January 20, 2022

Two-Sample T-test of Difference in Means

Supposed that we are interested in studying the relation between habits and practices of expectant mothers and the birth of their children. Using `ncbirths` dataset, our primary interest is the relationship between the mother's smoking status and the baby's weight. We construct the hypothesis test, and use `t.test()` to examine whether the mean difference between the nonsmoker and smoker mothers is statistically significant, and calculate confidence interval for the difference between the two.

$$H_0 : \mu_{\text{nonsmoker}} - \mu_{\text{smoker}} = 0$$
$$H_A : \mu_{\text{nonsmoker}} - \mu_{\text{smoker}} \neq 0$$

Base R

In base R, we can find the observed test statistics, p-value and confidence interval of the difference in means by passing the variables as **formula** argument into `t.test()`.

```
t.test(weight ~ habit, data=ncbirths)
```

```
##
##  Welch Two Sample t-test
##
## data:  weight by habit
## t = 2.359, df = 171.32, p-value = 0.01945
## alternative hypothesis: true difference in means between group nonsmoker and group smoker is not equa
## 95 percent confidence interval:
```

```
##  0.05151165 0.57957328
## sample estimates:
## mean in group nonsmoker    mean in group smoker
##                7.144273                6.828730
```

Now suppose that we are interested incorporating permutation on null hypothesis.

```r
# do permutation test here
```

infer

In `infer` package provides a series of functions that provides more clarity to the process. Under the hood, these functions are wrapper functions to the base R functions we used in the previous section.

Null distribution:

```r
set.seed(42)
null <- ncbirths %>%
  specify(weight ~ habit) %>%
  hypothesize(null = "independence") %>%
  generate(reps = 1000, type = "permute") %>%
  calculate(stat = "diff in means", order = c("nonsmoker", "smoker"))
```

```r
observed <- ncbirths %>%
  specify(weight ~ habit) %>%
  calculate(stat = "diff in means", order = c("nonsmoker", "smoker"))
```

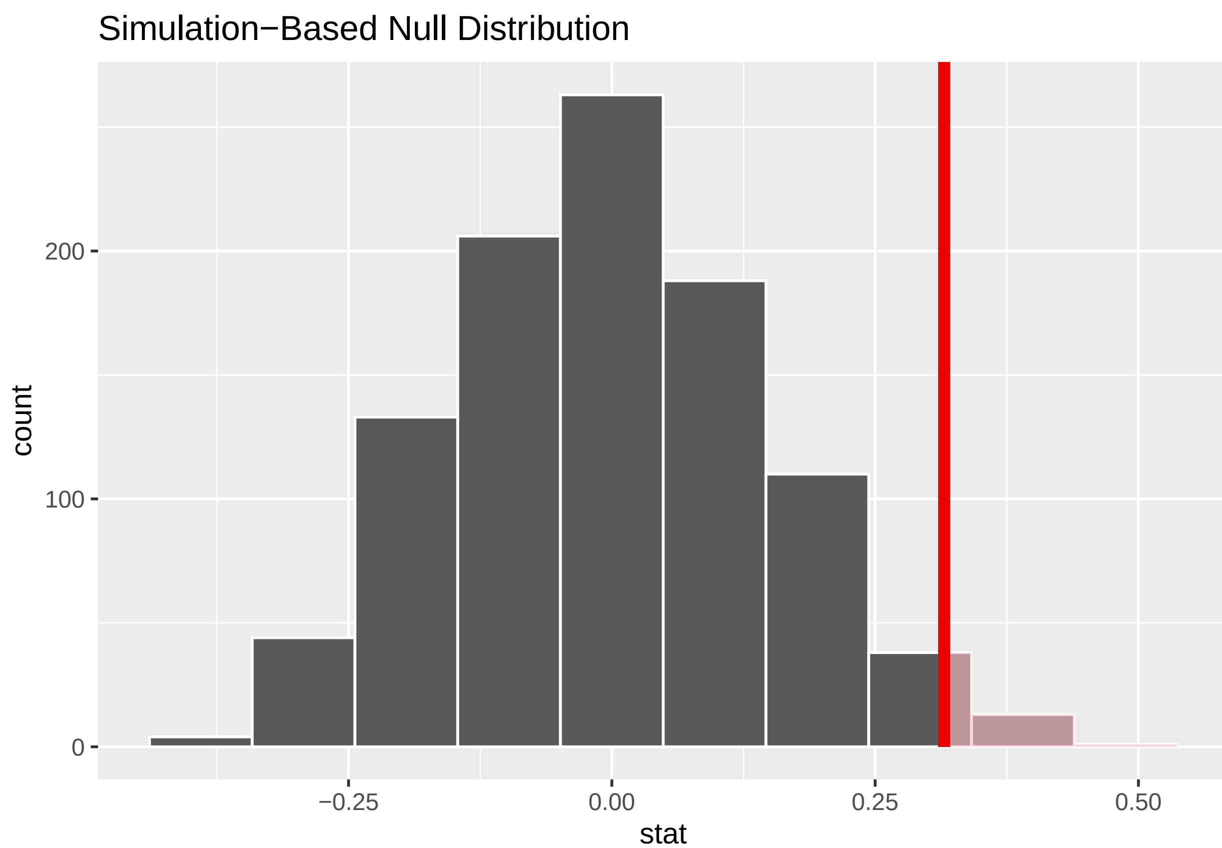
p-value:

```r
null %>%
  get_p_value(obs_stat = observed, direction = "right")
```

```
## # A tibble: 1 x 1
##   p_value
##     <dbl>
## 1   0.022
```

Visualization p-value

```r
visualize(null, bins = 10) +
  shade_p_value(obs_stat = observed, direction = "right")
```



Or using wrapper function

```
t_test(x = ncbirths,
       formula = weight ~ habit,
       order = c("nonsmoker", "smoker"),
       alternative = "two-sided")
```

```
## # A tibble: 1 x 7
##   statistic  t_df p_value alternative estimate lower_ci upper_ci
##       <dbl> <dbl>   <dbl> <chr>          <dbl>    <dbl>    <dbl>
## 1      2.36  171.  0.0195 two.sided      0.316   0.0515    0.580
```

Chi-Square Test of Independence

With the the same dataset, suppose that we are and we are interested in examining the dependence.

data arguemnt, and formule inferace

Chi-sq Test