

Wald test calibration

Kaitlyn Lagattuta

2024-11-22

Here, we use permutation to check the calibration of Wald test when applying mixed effects logistic regression to our testing data.

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
library(broom.mixed)
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.2
```

```
check_calibration <- function(data){
  set.seed(1)

  it = 1000
  p = numeric(length=it)
  drs = unique(data$Donor)

  for (i in 1:it){
    for (j in 1:length(drs)){
      data$TCRscore[data$Donor==drs[j]] = sample(data$TCRscore[data$Donor==drs[j]], replace=FALSE)
    }
    fit.perm = glmer(target ~ TCRscore + (1|Donor), data=data, family="binomial")
    p[i] = tidy(fit.perm)$p.value[2]
  }

  df = data.frame(expected = -log10(sort(runif(length(p)))), observed = -log10(sort(p)))
  g = ggplot(df, aes(x=expected, y=observed))
  g = g + geom_abline() + geom_smooth(method="lm", color="gray") + geom_point() + theme_bw() + xlab("-log10(p-value)")

  print("Type 1 error rate at alpha=0.01:")
  print(length(p[p<0.01])/length(p))
  return(g)
}
```

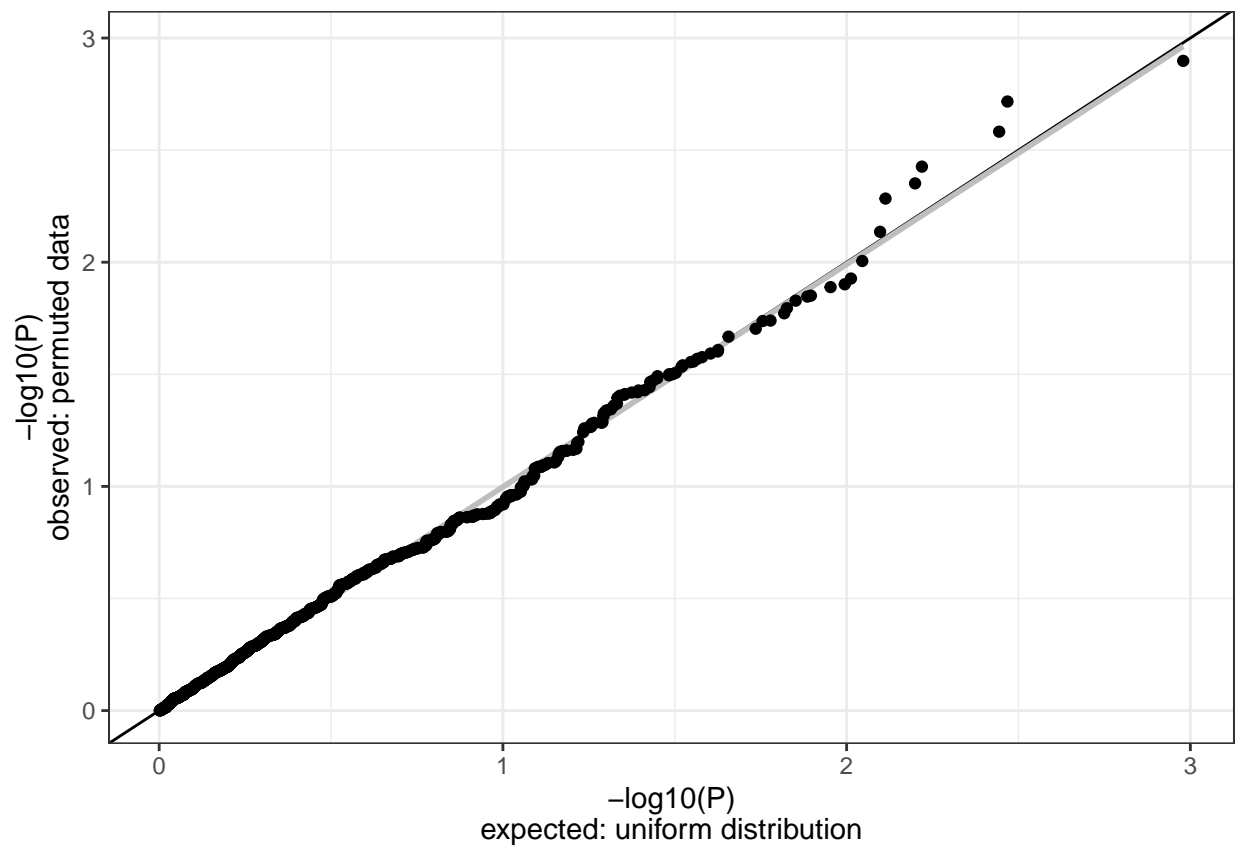
```
data = readRDS("data/Dataset3_scored_wannotations.rds")
```

Rare binomial outcome: PLZF-high innate-like T cells

```
data$TCRscore = data$TCR.innate
data$target = data$target1
check_calibration(data)
```

```
## [1] "Type 1 error rate at alpha=0.01:"
## [1] 0.008
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



Common binomial outcome: memory T cells

```
data$TCRscore = data$TCR.mem
data$target = data$target4
data = data[!(is.na(data$target)),] ##filters out innate-like T cells
check_calibration(data)
```

```
## [1] "Type 1 error rate at alpha=0.01:"
## [1] 0.006
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

