

Stat 344 – PS04

Luca Seazzu

June 01, 2023

Problem 5.18

```
x <- 35
n <- 55

LL <- function(L0, x, n){
  odds <- exp(L0)
  pi <- odds / (1 + odds)
  x * log(pi) + (n-x) * log(1-pi)
}

maxLik(LL, start = 0.5, x = x, n = n)

## Maximum Likelihood estimation
## Newton-Raphson maximisation, 3 iterations
## Return code 1: gradient close to zero (gradtol)
## Log-Likelihood: -36.0515 (1 free parameter(s))
## Estimate(s): 0.5596158
L0.hat <- 0.5596158

SE <- sqrt(L0.hat * (1-L0.hat) / n)

L0.hat + c(-1, 1) * qnorm(0.975) * SE

## [1] 0.4284177 0.6908139
Wald Interval is from 0.4284177 to 0.6908139
pval_minus_critical <- function(L0.0) {
  2 * (LL(L0.hat, x, n) - LL(L0.0, x, n)) - qchisq(.95, df = 1)
}

lo <- uniroot(pval_minus_critical, c(-100, L0.hat)) |> value()
hi <- uniroot(pval_minus_critical, c(L0.hat, 100)) |> value()

## Warning in uniroot(pval_minus_critical, c(L0.hat, 100)): NA/Inf replaced by
## maximum positive value
c(lo,hi)

## [1] 0.02050246 1.12635375
Likelihood Confidence Interval from 0.02050246 to 1.12635375.
```

Problem 5.48

a.

For large values of T we would reject the null hypothesis because the more the value gets away from 1, the less like a Poisson distribution it would look like. Not too sure about really small values since that would mean a really small sample mean with a really high sample variance.

b.

```
x <- c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 2, 2, 2, 4)

m <- mean(x)
s <- var(x)
n <- length(x)
T <- (n-1) * s/m

p <- 1- pchisq(T, (n-1))
p
```

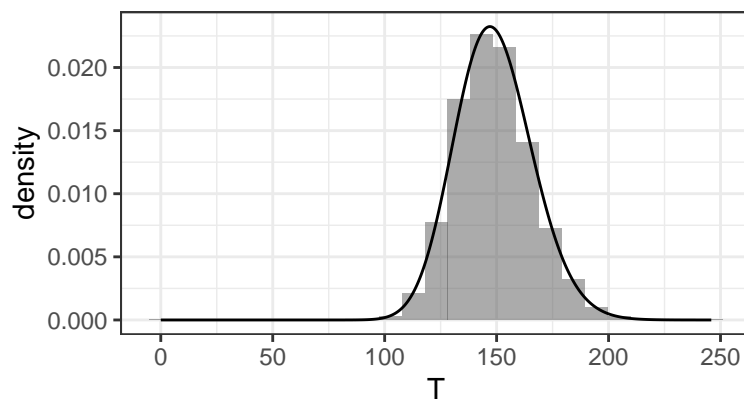
```
## [1] 0.02994042
```

A p-value of 0.029 would reject the null hypothesis that the distribution is Poisson so this would suggest x is not a Poisson distribution.

c.

```
T <- function(x){
  m <- mean(x)
  s <- var(x)
  n <- length(x)
  T <- (n-1) * s/m
}

sim <- do(10000) * T(rpois(150, lambda = 1))
gf_dhistogram( ~T, data = sim) |>
  gf_dist(dist = 'chisq', df = 149)
```



Simulating data with a larger sample size shows the quality of the chi-squared approximation and that the larger the sample size, the better the results.

Non-book Problems

```
gapminder <- read.csv('https://sldr.netlify.com/data/gapminder_clean.csv') |>
  na.omit()
```

1.

```
LL2 <- function(lambda, x, n) {
  if (lambda < 0) return(NA)
  sum(dexp(x, rate = lambda, log = TRUE))
}
```

```
maxLik(LL2, start = .5, x = gapminder$income)
```

```
## Maximum Likelihood estimation
## Newton-Raphson maximisation, 9 iterations
## Return code 8: successive function values within relative tolerance limit (reltol)
## Log-Likelihood: -38282.38 (1 free parameter(s))
## Estimate(s): 0.0001646964
```

MLE = 0.0001646964

2.

```
lambda.hat <- 0.0001646964
n2 <- length(gapminder$income)
```

```
SE <- lambda.hat / sqrt(n2)
```

```
lambda.hat + c(-1, 1) * qnorm(0.975) * SE
```

```
## [1] 0.0001595551 0.0001698377
```

Wald interval is from 0.0001595551 to 0.0001698377

```
pval_minus_critical2 <- function(lambda0) {
  2 * (LL2(lambda.hat, x, n) - LL2(lambda0, x, n)) - qchisq(.95, df = 1)
}
```

```
lo2 <- uniroot(pval_minus_critical2, c(0, lambda.hat)) |> value()
hi2 <- uniroot(pval_minus_critical2, c(lambda.hat, 100)) |> value()
## Question
c(lo2, hi2)
```

```
## [1] 1.646964e-04 1.438195e+01
```

Likelihood-ratio confidence interval is from .0001646964 to 14.38

3.

```
llgamma <- function(alpha, var, x){
  if (alpha < 0) return(NA)
  if (var < 0) return(NA)
  log(dgamma(x, shape = alpha, scale = var))
}
```

```
log_cost <- log(gapminder$income)
```

```
ll_alternate <- maxLik(logLik = llgamma,
```

```
start = c(0.1, 1),  
x = log_cost)
```

```
logLik(ll_alternate) |> as.numeric()
```

Couldn't quite figure problem #3...