# list1

Kamil Zaborniak

## Contents

# Task 1

We have the simple random sample: $\underline{X_n} = (X_1, \ldots, X_n)$ from the Beta distribution: $Beta(\alpha+1, 1)$ with the density:

$$f(x, \alpha) = \frac{\Gamma(\alpha+1+1)}{\Gamma(\alpha+1)\Gamma(1)} x^{\alpha+1-1}(1-x)^{1-1} = (\alpha+1)x^{\alpha}$$

for $x \in (0, 1)$ and $\alpha > -1$.

## MLE

To find the maximum likelihood estimator, let's start by determining the likelihood function for sample $\underline{X_n}$:

$$L\left(\underline{X_n}, \alpha\right) = \prod_{i=1}^{n} f\left(X_i, \alpha\right) = \prod_{i=1}^{n} (\alpha+1) X_i^{\alpha} = (\alpha+1)^n \prod_{i=1}^{n} X_i^{\alpha}.$$

In the second step, let us determine the log-likelihood function, which is a transformation of the above function and has the form:

$$l\left(\underline{X_n}, \alpha\right) = \log\left(L\left(\underline{X_n}, \alpha\right)\right) = \log\left((\alpha+1)^n \prod_{i=1}^{n} X_i^{\alpha}\right) = n\log(\alpha+1) + \alpha \sum_{i=1}^{n} \log\left(X_i\right).$$

To find the maximum likelihood estimator, we need to find the value for which the value of the likelihood function is the largest. The argument that maximizes the log-likelihood function is also the argument that maximizes the log-likelihood function, but finding it is easier for the log-likelihood function. To determine the desired estimator, examine the variability of the log-likelihood function with respect to $\alpha$. Let's determine the first derivative of the function and its zeros:

$$\frac{\partial l\left(\underline{X_n}, \alpha\right)}{\partial \alpha} = \frac{n}{\alpha+1} + \sum_{i=1}^{n} \log\left(X_i\right),$$

$$\frac{n}{\alpha+1} + \sum_{i=1}^{n} \log\left(X_i\right) = 0 \iff \alpha = \frac{-n}{\sum_{i=1}^{n} \log\left(X_i\right)} - 1.$$

Let's determine the second derivative and examine its sign:

$$\frac{\partial^2 l\left(\underline{X_n}, \alpha\right)}{\partial \alpha^2} = \frac{-n}{(\alpha+1)^2}.$$

The second derivative of the log-likelihood function is negative for all $\alpha > -1$, so the function is decreasing. Consequently, the maximum argument of log-likelihood function and the maximum likelihood estimator estimator is:

$$\hat{\alpha}_{MLE} = \frac{-n}{\sum_{i=1}^{n} \log\left(X_i\right)} - 1.$$

## Fisher Information

In order to calulate the Fisher Information, let us start by determining the first and second derivatives of logarithm of the density $\log\left(f(x, \alpha)\right) = \log(\alpha+1) + \alpha\log(x)$:

$$\frac{\partial}{\partial \alpha}\left(\log(\alpha+1) + \alpha\log(x)\right) = \log(x) + \frac{1}{\alpha+1}.$$

$$\frac{\partial^2}{\partial \alpha^2}\left(\log(\alpha+1) + \alpha\log(x)\right) = \frac{-1}{(\alpha+1)^2}.$$

Let's determine the Fisher information:

$$I(\alpha) = -\mathbb{E}\left[\frac{\partial^2 f(x,\alpha)}{\partial \alpha^2}\right] = \frac{1}{(\alpha+1)^2}.$$

## Distribution of $\hat{\alpha}_{MLE}$

If $X_1, \ldots, X_n$ are i.i.d with pdf: $f(x,\alpha)$ and $0 < I(\alpha) < \infty$, then we know:

$$\sqrt{n}\left(\hat{\alpha}_{MLE} - \alpha\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{1}{I(\alpha)}\right).$$

Consequently:

$$\hat{\alpha}_{MLE} \sim \mathcal{N}\left(\alpha, \frac{(\alpha+1)^2}{n}\right).$$

## Bias of $\hat{\alpha}_{MLE}$

Maximum likelihood estimator is unbiased, so:

$$b\left[\hat{\alpha}_{MLE}\right] = \mathbb{E}\left[\alpha - \hat{\alpha}_{MLE}\right] = \alpha - \mathbb{E}\left[\hat{\alpha}_{MLE}\right] = 0.$$

## Mean Squared Error of $\hat{\alpha}_{MLE}$

Based on the bias and variance of the maximum likelihood estimator $\hat{\alpha}_{MLE}$, let's calculate its mean squared error:

$$MSE\left[\hat{\alpha}_{MLE}\right] = Var\left[\hat{\alpha}_{MLE}\right] + \left(b\left[\hat{\alpha}_{MLE}\right]\right)^2 = \frac{(\alpha+1)^2}{n}.$$

## Moment estimator of $\alpha$

The idea behind Method of Moments estimation is to find a good estimator. For this purpose we should have the true and sample moments match as best we can. That is, I should choose the parameter $\alpha$ such that the first true moment is equal to the first sample moment .

Start by determining the first moment of the $Beta(\alpha+1, 1)$ distribution:

$$\mathbb{E}[X] = \int_0^1 (\alpha+1)x^\alpha \cdot x\, dx = \frac{\alpha+1}{\alpha+2} x^{\alpha+2}\Big|_0^1 = \frac{\alpha+1}{\alpha+2}.$$

The formula of the first moment of the sample $\underline{x} = (x_1, \ldots, x_n) \sim Beta(\alpha+1, 1)$ is:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Hence we have:

$$\overline{x} = \frac{\alpha+1}{\alpha+2} \Longrightarrow \hat{\alpha}_{MOM} = \frac{2x-1}{1-x}.$$

## Simulations

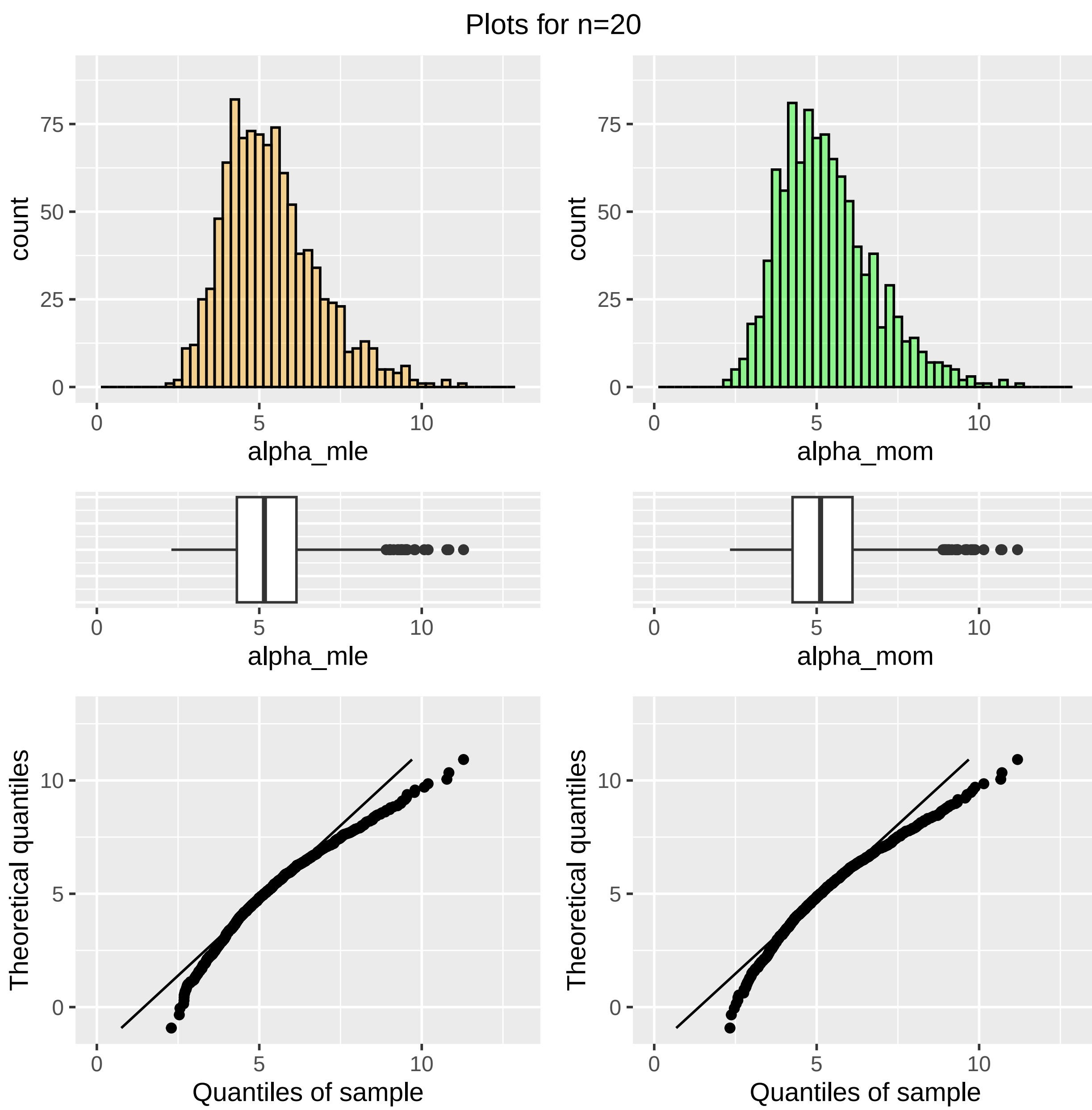
Let's fix $\alpha = 5$ and set seed to 411. Then generate sample from $Beta(6, 1)$ of the size $n = 20$.

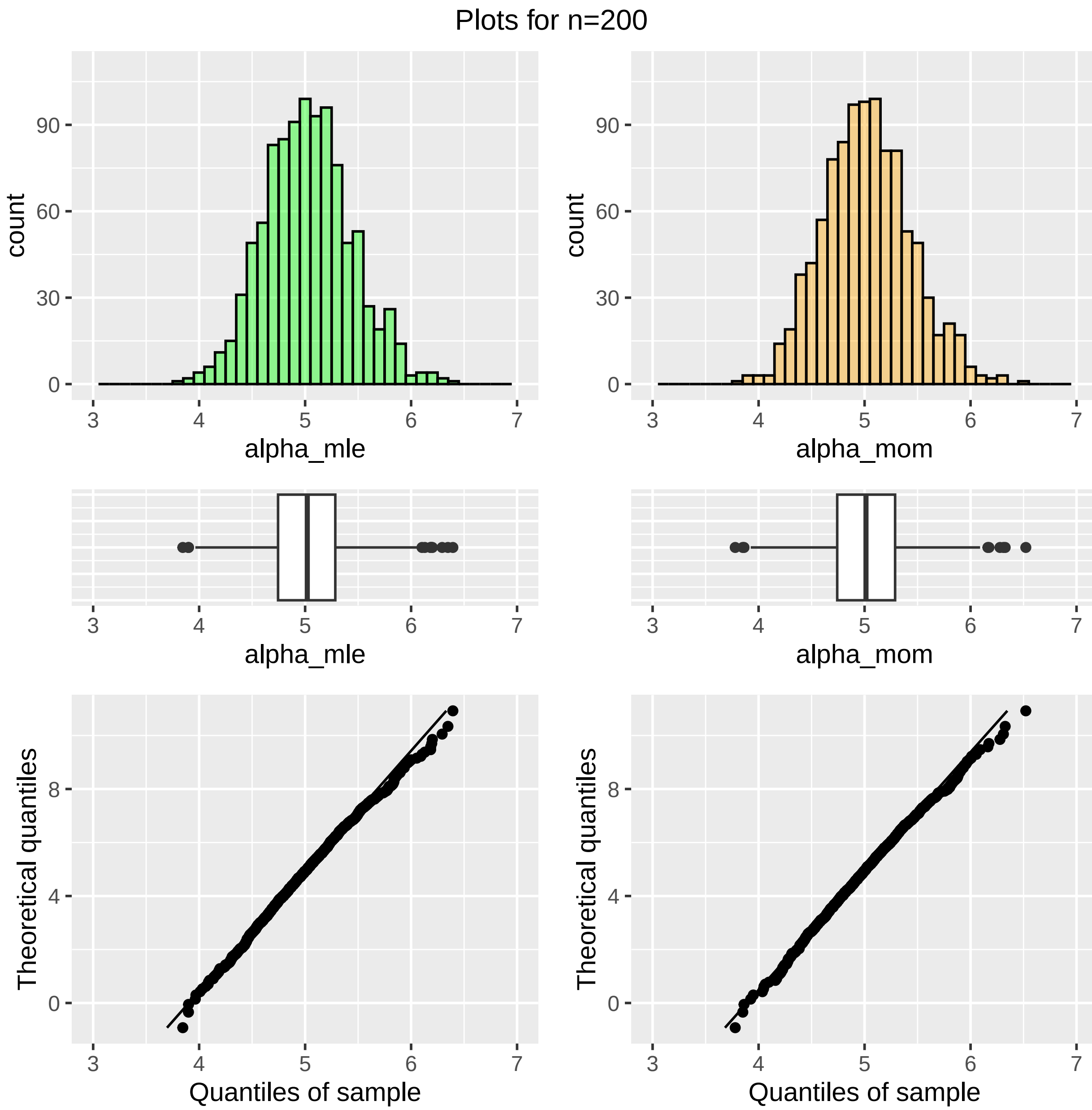
As a result of the simulation, we obtained the following results:

| Estimator | Value | $\alpha - \hat{\alpha}$ | $(\alpha - \hat{\alpha})^2$ |
|---|---|---|---|
| $\hat{\alpha}_{MLE}$ | 5.9515 | -0.9515 | 0.9053 |
| $\hat{\alpha}_{MOM}$ | 5.7600 | -0.7600 | 0.5776 |

A better estimator is the moment estimator. The value of $(\alpha - \hat{\alpha})^2$ is smaller, but the differences between the estimators are too small to determine which is better. For this purpose, we will repeat the simulation 1000 times.

First, we will generate 1000 n-element(n=20 and n=200) samples from the $Beta(6, 1)$ distribution. Then, for each sample, we will calculate the maximum likelihood estimator and the moment's estimator. We will draw a histogram, boxplot and qqplot for the resulting samples of estimators.


Plots for n=20
count
75
50
25
0
0
5
10
alpha_mle
alpha_mom
Theoretical quantiles
Quantiles of sample




We can see in the plots above that the distributions of the maximum-likelihood estimators and moment estimators are similar to the normal distribution. In the case when n=20, the distribution of the studied estimators is close to the distribution of $N(5, 1.8)$, in the case when n=200 - $N(5, 0.18)$. The plots confirm that the maximum-likelihood estimators and moment estimators asymptotically come from the normal distribution. For sample size n=200, we see that the distribution of estimates is more symmetrical and more 'normal' compared to the case when n=20. Moreover, we also see that as the sample size increases, the variance of the estimates decreases, which is true because it is inversely proportional to the sample size (n). However, regardless of the sample size, the expected value of the estimators remains equal to 5. When comparing maximum-likelihood estimators and moment estimators, we cannot detect significant differences between them regardless of sample size. In my opinion, based solely on the above charts, the probability distributions of the studied estimators are very similar. the bias, the variance and the mean-squared error of both estimators and construct approximate 95% confidence intervals for these parameters.

To further examine the differences between the maximum likelihood estimator and the moments estimator, based on simulation samples, let's estimate the bias, the variance and the mean-squared error of both estimators and construct approximate 95% confidence intervals for these parameters.

### Bias

The bias of the $\hat{\alpha}$ estimator, based on the sample of estimators: $(\hat{\alpha}_1, \ldots, \hat{\alpha}_{1000})$, is calculated as follows:

$$b(\hat{\alpha}) = 5 - \frac{1}{1000}\sum_{i=1}^{1000} \hat{\alpha}_i.$$

We know that:

$$\frac{b(\hat{\alpha}) - (5 - \hat{\alpha}_i)}{\sigma(5 - \hat{\alpha}_i)}\sqrt{1000} \sim N(0,1),$$

where $\sigma(\cdot)$ is the standard deviation. Hence, the 95% confidence interval of the bias is:

$$b(\hat{\alpha}) \pm z_{0.975}\frac{\sigma(5 - \hat{\alpha}_i)}{\sqrt{1000}},$$

where $z_{0.975}$ is the 97.5% quantile of standard normal distribution.

### Variance

The variance of the $\hat{\alpha}$ estimator, based on the sample of estimators: $(\hat{\alpha}_1, \ldots, \hat{\alpha}_{1000})$, is calculated as follows:

$$\sigma^2(\hat{\alpha}) = \frac{1}{999}\sum_{i=1}^{1000} (\hat{\alpha}_i - 5)^2.$$

We know that:

$$\frac{999\sigma^2(\hat{\alpha}_i)}{\sigma^2(\hat{\alpha})} \sim \chi^2_{999},$$

so the 95% confidence interval of the variance is:

$$\left[\frac{999\sigma^2(\hat{\alpha}_i)}{\chi^2_{0.975;999}}, \frac{999\sigma^2(\hat{\alpha}_i)}{\chi^2_{0.025;999}}\right].$$

Where $\chi^2_{p,999}$ is the $p$ quantile of the $\chi^2$ distribution with 999 df.

From the Cramer-Rao's inequality, we know that the variance of the maximum-likelihood estimator is:

$$\sigma^2(\hat{\alpha}) \geq \frac{36}{n}.$$

### MSE

Similarly, the mean square error of the $\hat{\alpha}$ estimator, based on the sample of estimators:

$$MSE(\hat{\alpha}) = \frac{1}{1000}\sum_{i=1}^{1000}(5 - \hat{\alpha}_i)^2 = \sigma^2(\hat{\alpha}) + (b(\hat{\alpha}))^2$$

and:

$$\frac{MSE(\hat{\alpha}) - (5 - \hat{\alpha}_i)^2}{\sigma((5 - \hat{\alpha}_i)^2)} \sim N(0,1).$$

The 95% confidence interval of the MSE is:

$$MSE(\hat{\alpha}) \pm z_{0.975}\frac{\sigma((5 - \hat{\alpha}_i)^2)}{\sqrt{1000}}.$$

**Results**

Based on the experiment, we obtained the following values:

| | n=20 | | n=200 | |
|---|---|---|---|---|
| | $\hat{\alpha}_{MLE}$ | $\hat{\alpha}_{MOM}$ | $\hat{\alpha}_{MLE}$ | $\hat{\alpha}_{MOM}$ |
| Theoretical bias | 0 | - | 0 | - |
| Bias | 0.358 | 0.315 | 0.028 | 0.024 |
| 95% confidence intervals of bias | $[0.269, 0.447]$ | $[0.226, 0.404]$ | $[0.003, 0.054]$ | $[-0.002, 0.050]$ |
| Theoretical variance | $\geq 1.8$ | - | $\geq 0.18$ | - |
| Variance | 2.197 | 2.182 | 0.169 | 0.171 |
| 95% confidence intervals of variance | $[2.017, 2.404]$ | $[2.003, 2.387]$ | $[0.155, 0.185]$ | $[0.157, 0.187]$ |
| Theoretical MSE | 1.8 | - | 0.18 | - |
| MSE | 2.325 | 2.282 | 0.169 | 0.172 |
| 95% confidence intervals of MSE | $[2.079, 2.572]$ | $[2.041, 2.522]$ | $[0.155, 0.185]$ | $[0.156, 0.187]$ |

The obtained results confirm that as the sample size increases, the biases of both estimators tend to 0, their variances decrease and are closer to $\frac{36}{n}$, and their mean square error decreases. Regardless of the sample size, it is difficult to indicate a significant difference between the maximum likelihood estimator and the moments estimator. Both are unbiased, have similarly low variance and mean square error. Taking into account the results when n=20, the better estimator is the moments estimator, in the case when n=200 the better estimator is the maximum likelihood estimator (although its bias is higher, its mean square error is smaller).

# Task 2

Let $\underline{X_n} = (X_1, \ldots, X_n)$ be the simple random sample from the distribution with the density $f(x) = \lambda e^{-\lambda x}$, for $x > 0$, $\lambda > 0$. Using the Neyman-Pearson lemma, let us test at the level $\alpha = 0.05$:

$$H_0 : \lambda = 5,$$

$$H_1 : \lambda = 3.$$

Let's start by determining the likelihood ratio:

$$R = \frac{\prod_{i=1}^n f_0(X_i)}{\prod_{i=1}^n f_1(X_i)} = \frac{\prod_{i=1}^n 5e^{-5X_i}}{\prod_{i=1}^n 3e^{-3X_i}} = \left(\frac{5}{3}\right)^n \prod_{i=1}^n e^{-2X_i} = \left(\frac{5}{3}\right)^n \exp\left\{-2\sum_{i=1}^n X_i\right\}.$$

## Critical Value

We reject $H_0$, when:

$$
\begin{aligned}
R &\leq k, \\
\left(\frac{5}{3}\right)^n \exp\left\{-2\sum_{i=1}^{n} X_i\right\} &\leq k, \\
\exp\left\{-2\sum_{i=1}^{n} X_i\right\} &\leq \left(\frac{3}{5}\right)^n k, \\
-2\sum_{i=1}^{n} X_i &\leq n\log\left(\frac{3}{5}\right) + \log\left(k\right), \\
\sum_{i=1}^{n} X_i &\geq \frac{-n\log\left(\frac{3}{5}\right) - \log\left(k\right)}{2} = k^*.
\end{aligned}
$$

We know that $\sum_{i=1}^{n} X_i \sim Gamma(n, \lambda)$, which means that the critical value is equal to:

$$k^* = F^{-1}_{Gamma(n,5)}(1-\alpha).$$

Where $F^{-1}_{Gamma(n,5)}(1-\alpha)$ is the $1-\alpha$ quantile of the $Gamma(n, 5)$ distribution.

## Power of test

We know that under $H_1$:

$$\sum_{i=1}^{n} X_i \sim Gamma(n, 3),$$

so, the power of the test is:

$$\gamma(\underline{X_n}, \alpha) = P_{H_1}\left(\sum_{i=1}^{n} X_i \geq k^*\right) = 1 - F_{Gamma(n,3)}\left(F^{-1}_{Gamma(n,5)}(1-\alpha)\right).$$

What is more, when $n \to \infty$:

$$\gamma(\underline{X_n}, \alpha) \to 1$$

## p-value

Taking the test statistic $T\left(\underline{X_n}\right) = \sum_{i=1}^{n} X_i$, we can use it to determine the p-value. Suppose $T \sim\sim Gamma(n, 5)$, then:

$$p = P_{H_0}\left(T\left(\underline{X_n}\right) < T\right) = 1 - F_{Gamma(n,5)}\left(T\left(\underline{X_n}\right)\right).$$

Suppose $n = 20$, and generate n-size random samples from exponential distribution under $H_0$ and under $H_1$. then under $H_0$: $p = 0.1477$ and under $H_1$: $p = 9.518 \cdot 10^{-5}$. This confirms that, under $H_1$, the p-value takes extremely low values.

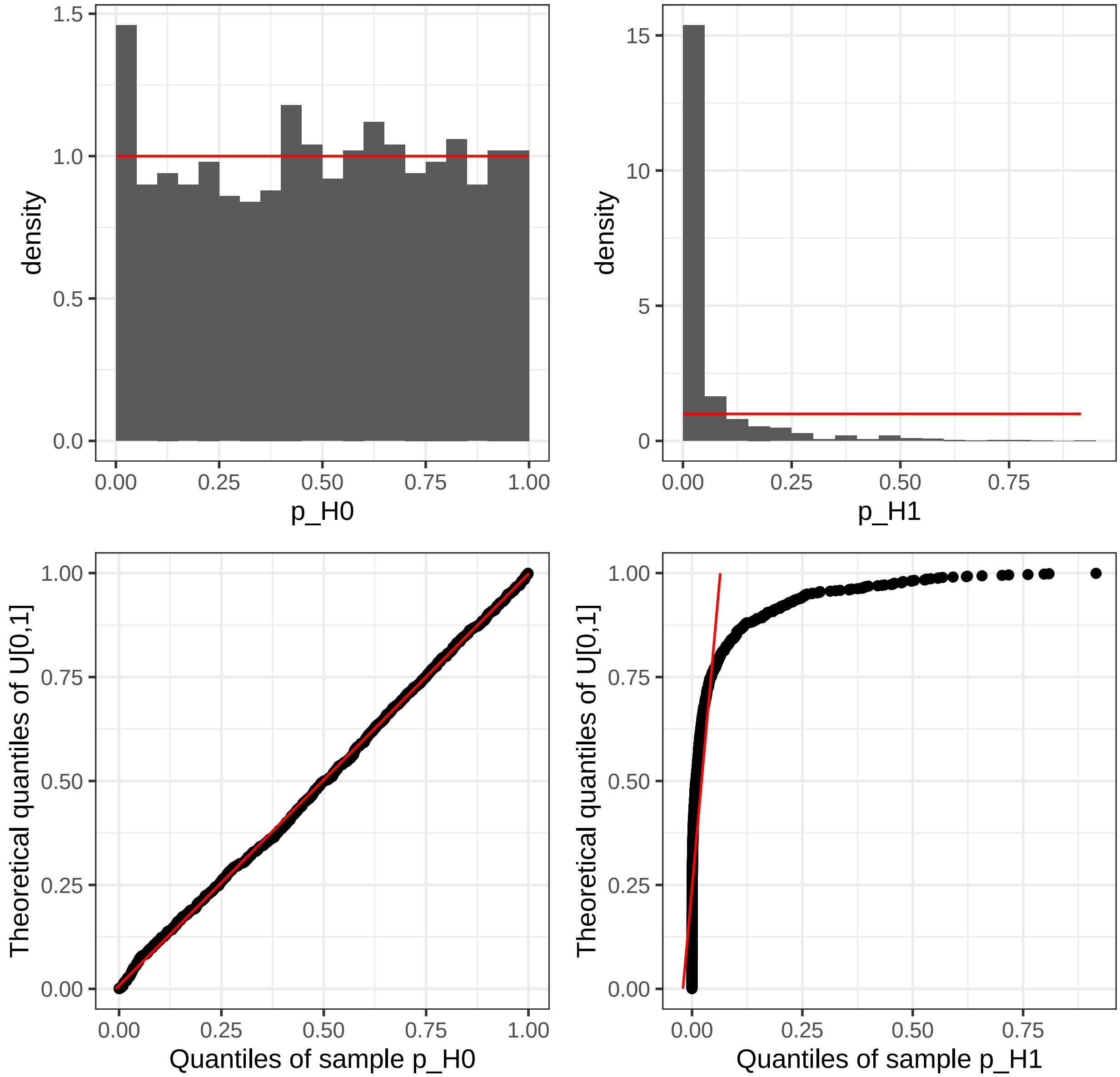
Under $H_0$, we know that p-value is uniformly distributed on $[0,1]$, if the test statistic comes from a continuous distribution. Let's prove it with our example, with $x \in [0,1]$:

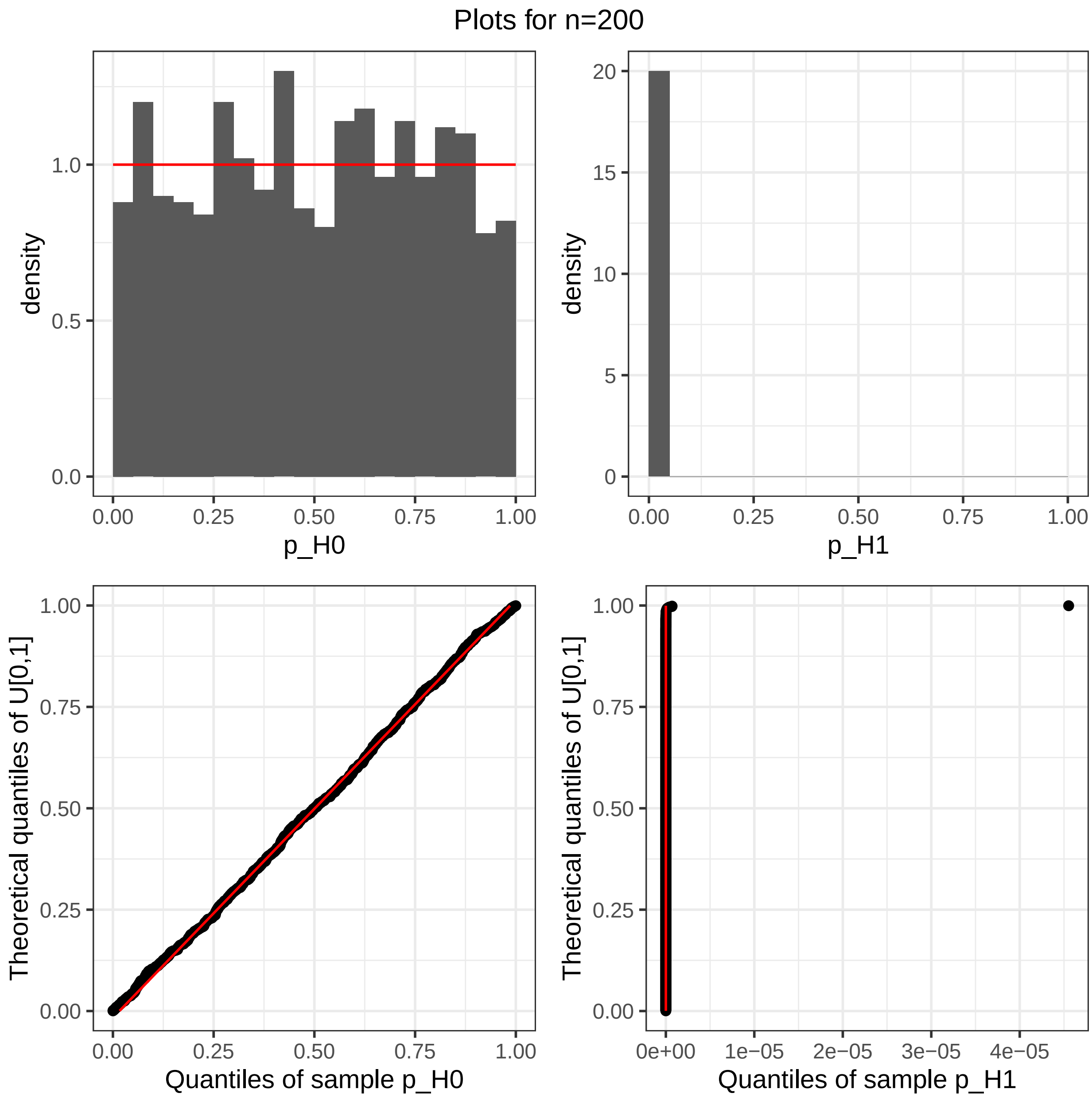
$$\begin{aligned}
P_{H_0}(p < x) &= P_{H_0}\left(1 - F_{Gamma(n,5)}\left(T\left(\underline{X_n}\right)\right) < x\right) \\
&= P_{H_0}\left(F_{Gamma(n,5)}\left(T\left(\underline{X_n}\right)\right) > 1 - x\right) \\
&= P_{H_0}\left(T\left(\underline{X_n}\right) > F^{-1}_{Gamma(n,5)}(1-x)\right) \\
&= 1 - F_{Gamma(n,5)}\left(F^{-1}_{Gamma(n,5)}(1-x)\right) \\
&= 1 - 1 + x \\
&= x.
\end{aligned}$$

## Simulation

We will generate 1000 n-element samples (for n=20 and n=200) from the exponential distribution under H0 and under H1. Then, for each case, let's determine p-values based on the obtained samples. Let's analyze their histogram and qqplots.

Plots for n=20






We see that under $H_0$, the p-values are uniformly distributed( qqplots show it more clearly, the point are arranged along the x=y line) and does not depend on the sample size. Under $H_1$, as the size of a single sample increases, they take on more and more extremaly low values (they certainly do not assume a uniform distribution on [0,1]).

**Type I error**

Based on the simulation, let's try to estimate the probability of committing a type I error and its 95% confidence interval. Let $\underline{p}_{H_0} = \left(p_1^0, \ldots, p_n^0\right)$, ten:

$$\hat{P_{H_0}}\left(TypeIerror\right) = \frac{1}{1000}\sum_{i=1}^{1000}\mathbb{I}_{\{p_i^0 \leq \alpha\}} = \beta.$$

The 95% confidence interval has form:

$$\beta \pm z_{1-\frac{\alpha}{2}} \frac{\sqrt{\beta(1-\beta)}}{\sqrt{1000}}.$$

For $n = 20$, the estimated probability of Type I Error(based on samples generated under $H_0$) is equal to 0.073 and his 95% confidence interval has form: $[0.057, 0.089]$.

For $n = 200$,the estimated probability of Type I Error(based on samples generated under $H_0$) is equal to 0.044 and his 95% confidence interval has form: $[0.031, 0.058]$.

The above results show that as the size of the generated sample increases, we can better control the type I error at a given alpha significance level. We also see that the size of a single sample affects the length of the confidence interval.

### Power

Let's also use simulation to determine the power of the test based on $\underline{p}_{H_1} = \left(p_1^1, \ldots, p_n^1\right)$. Let's estimate power of the test:

$$1 - \hat{P_{H_1}}\left(TypeIIerror\right) = \frac{1}{1000} \sum_{i=1}^{1000} \mathbb{I}_{\{p_i^1 \leq \alpha\}} = \hat{\gamma}.$$

The 95% confidence interval has form:

$$\hat{\gamma} \pm z_{1-\frac{\alpha}{2}} \frac{\sqrt{\hat{\gamma}(1-\hat{\gamma})}}{\sqrt{1000}}.$$

For $n = 20$ the power of the test is equal to 0.769 and 95% CI has form: $[0.743, 0.795]$. For $n = 200$ the power of the test is equal to 0.769 and 95% CI has form: $[1, 1]$.

The results confirm that the power of the test converges to 1, when $n \to \infty$. However, a sample size of 200 is already sufficient to reject the null hypothesis when it is false, with probability close to 1.