

QM8

Simulationsbasierte Inferenz (SBI):  
ein Hypothesentest

# Hypothesen testen: formaler Aufbau

$H_1$ : Es gibt einen Unterschied (etwas ist größer/kleiner etc.)

$H_0$ : Es gibt **keinen** Unterschied

Beispiele:

*Frage: Sind die Ankünfte am JFK pünktlicher?*

$H_1$ : Die Flüge am JFK haben weniger Verspätung

$H_0$ : Die Flüge am JFK haben **nicht weniger/ gleiche** Verspätung

*Frage: Hilft der Impfstoff?*

$H_1$ : Der Impfstoff erhöht die Immunität

$H_0$ : Der Impfstoff erhöht **nicht** die Immunität

*Hilft: Ist die Münze fair?*

$H_1$ : Die Münze ist gezinkt

$H_0$ : Die Münze ist **nicht** gezinkt

$\mu$ : unbekannter **Mittelwert** in der Gesamtheit( Population)

( $\pi$ : unbekannter **Anteil** in der Gesamtheit( Population))

$$H_1: \mu_{\text{JFK}} < \mu_{\text{non\_JFK}}$$

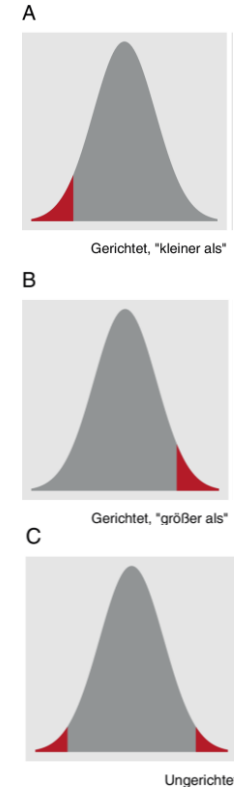
$$H_0: \mu_{\text{JFK}} \geq \mu_{\text{non\_JFK}}$$

$$H_1: \mu_{\text{impf}} > \mu_{\text{non\_impf}}$$

$$H_0: \mu_{\text{impf}} \leq \mu_{\text{non\_impf}}$$

$$H_1: \mu_{\text{Kopf}} \neq \mu_{\text{Zahl}}$$

$$H_0: \mu_{\text{Kopf}} = \mu_{\text{Zahl}}$$

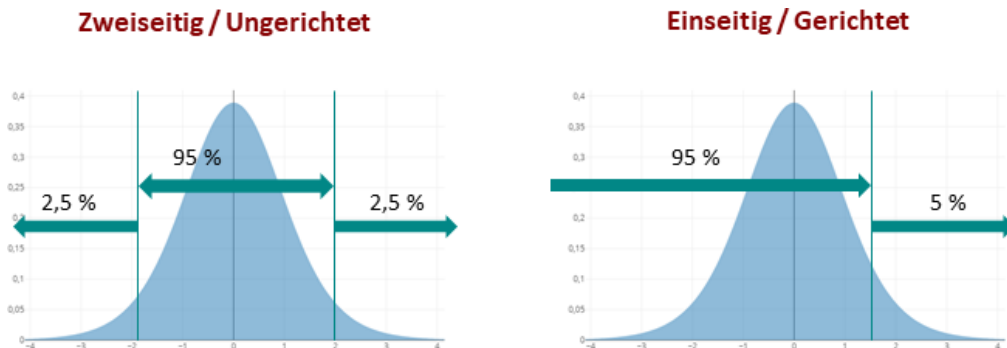


# gerichtete (einseitige) vs. ungerichtete (zweiseitige) Hypothesen

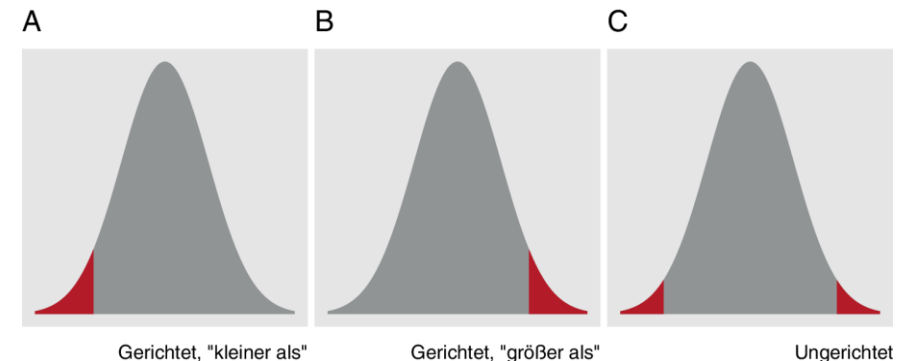
Ein Hypothesentest hat das Ziel eine wissenschaftliche Forschungsfrage mit empirischen Methoden zu beantworten. Dazu benötigen wir (in der Regel) Stichproben, die naturgemäß mit Unsicherheit behaftet sind (z.B. durch Stichprobenumfang und Repräsentativität).

Wir prüfen z.B. ob ein Mittelwert noch mit der Schwankung des Zufallsprozesses (Streuung) zu erklären ist oder nicht. Die Verteilung und Streuung der Stichprobenmittelwerte (Standardfehler) zeigt welche Werte unter der Nullhypothese (kein Unterschied) am Wahrscheinlichsten sind (basierend auf den vorliegenden Daten).

Die Übergänge zu den weniger wahrscheinlichen Beobachtungen sind fließend sodass eine Regel für die Entscheidung festgelegt werden muss: typischerweise sagen wir dass wenn eine Beobachtung eine Wahrscheinlichkeit von  $< 5\%$  hat, dieses Ergebnis *selten* ist. In diesem Fall wird die Grundannahme (dass es keine Auffälligkeit gibt: die Nullhypothese) dann verworfen (falsifiziert). Andernfalls, wenn das Ergebnis nicht selten ist unter der Grundannahme ( $> 5\%$ ) können wir die Nullhypothese nicht verwerfen. In diesem Fall wird sie bestätigt (verifiziert).



Beispiel: Intelligenzquotient

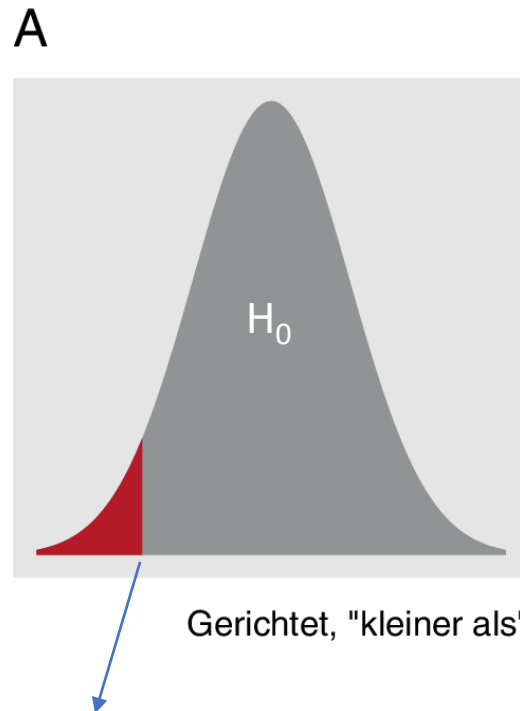


## Beispiel Hypothesentest

*Frage: Sind die Ankünfte am JFK weniger verspätet als an non\_JFK?*

$H_1$ : Die Flüge am JFK haben weniger Verspätung (Ankunft)  $\leftarrow$  gilt wenn  $< 5\%$

$H_0$ : Die Flüge am JFK haben **nicht weniger** Verspätung (Ankunft)  $\leftarrow$  bleibt bei  $\geq 5\%$



$H_1: \mu_{\text{JFK}} < \mu_{\text{non\_JFK}}$   $\leftarrow$  gilt wenn  $< 5\%$

$H_0: \mu_{\text{JFK}} \geq \mu_{\text{non\_JFK}}$   $\leftarrow$  bleibt bei  $\geq 5\%$

$\alpha = 5\%$  (auch 1% oder 0,1%)

Signifikanzniveau/ Irrtumswahrscheinlichkeit

```
QM6c QM7 QM8.R x flights x Funktionen.R x jfk x
Source on Save Run Source
56
57 ##### QM8 #####
58
59 # arr_delay filtern nach 'JFK' und 'non_JFK'
60 #flights_clean <- drop_na(flights, arr_delay) #ohne NA
61
62 # nur jfk
63 jfk <- flights_clean %>% filter(origin == "JFK") %>% select(arr_delay)
64 str(jfk) #109.079
65
66 # nicht jfk
67 non_jfk <- flights_clean %>% filter(origin != "JFK") %>% select(arr_de
68 str(non_jfk) #218.267
69 # 109.079 + 218.267 = 327.346
70
71 # Anzahl prüfen mit Pupulation (ohne NA)
72 str(flights_clean) #327,346 x 19 --> stimmt
73
```

JFK- Daten von den beiden anderen  
Flughäfen trennen (non\_jfk)

kontrollieren

```
Console Terminal x Jobs x
~/
> str(jfk) #109.079
tibble [109,079 x 1] (S3: tbl_df/tbl/data.frame)
 $ arr_delay: num [1:109079] 33 -18 -8 -2 -3 7 -4 -8 14 4 ...
> # nicht jfk
> non_jfk <- flights_clean %>% filter(origin != "JFK") %>% select(arr_delay)
> str(non_jfk) #218.267
tibble [218,267 x 1] (S3: tbl_df/tbl/data.frame)
```

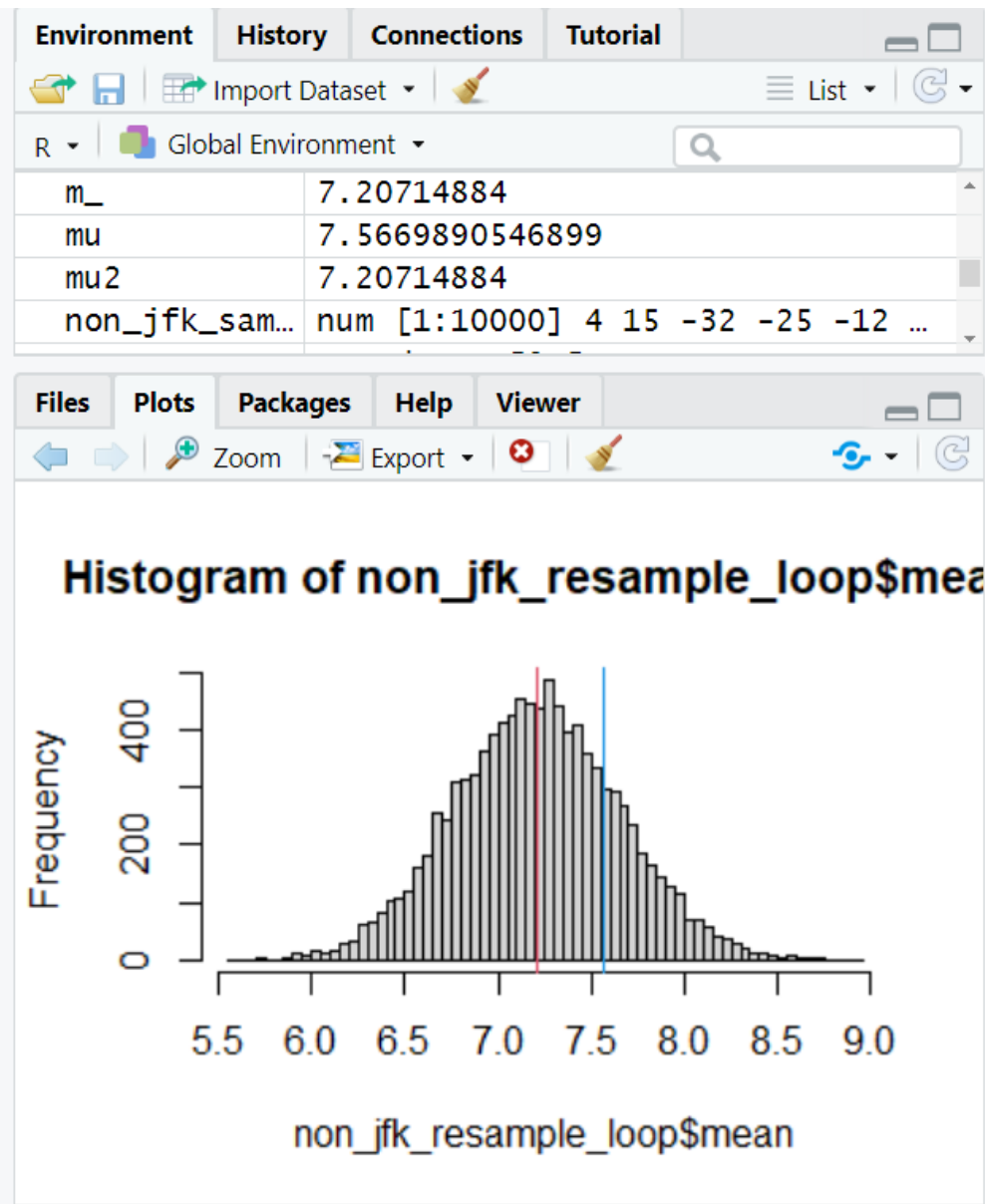
```
QM6c QM7 QM8.R x flights x Funktionen.R x jfk x
Source on Save Run Source
95 |
96 #####
97 ### H1: Die Flüge am JFK haben weniger Verspätung (Ankunft)
98 ### H0: Die Flüge am JFK haben nicht weniger Verspätung (Ankunft)
99
100
101 # ein JFK sample ziehen
102 jfk_sample<- sample(jfk$arr_delay, size = 10000)
103 x_<-mean(jfk_sample) #an dieser Stelle laesst sich nur der
104 # mean des samples berechnen
105 x_
106
107 non_jfk_sample<- sample(non_jfk$arr_delay, size = 10000)
108 mu<-mean(non_jfk_sample) #an dieser Stelle laesstt sich nur der
109 # mean des samples berechnen
110 mu
111
112
113
114
95:1 # QM8 R Script
Console Terminal x Jobs x
~/
> x_
[1] 6.0769
> mu
[1] 7.2063
```

Ein sample aus jfk ziehen

Ein sample aus non\_jfk ziehen

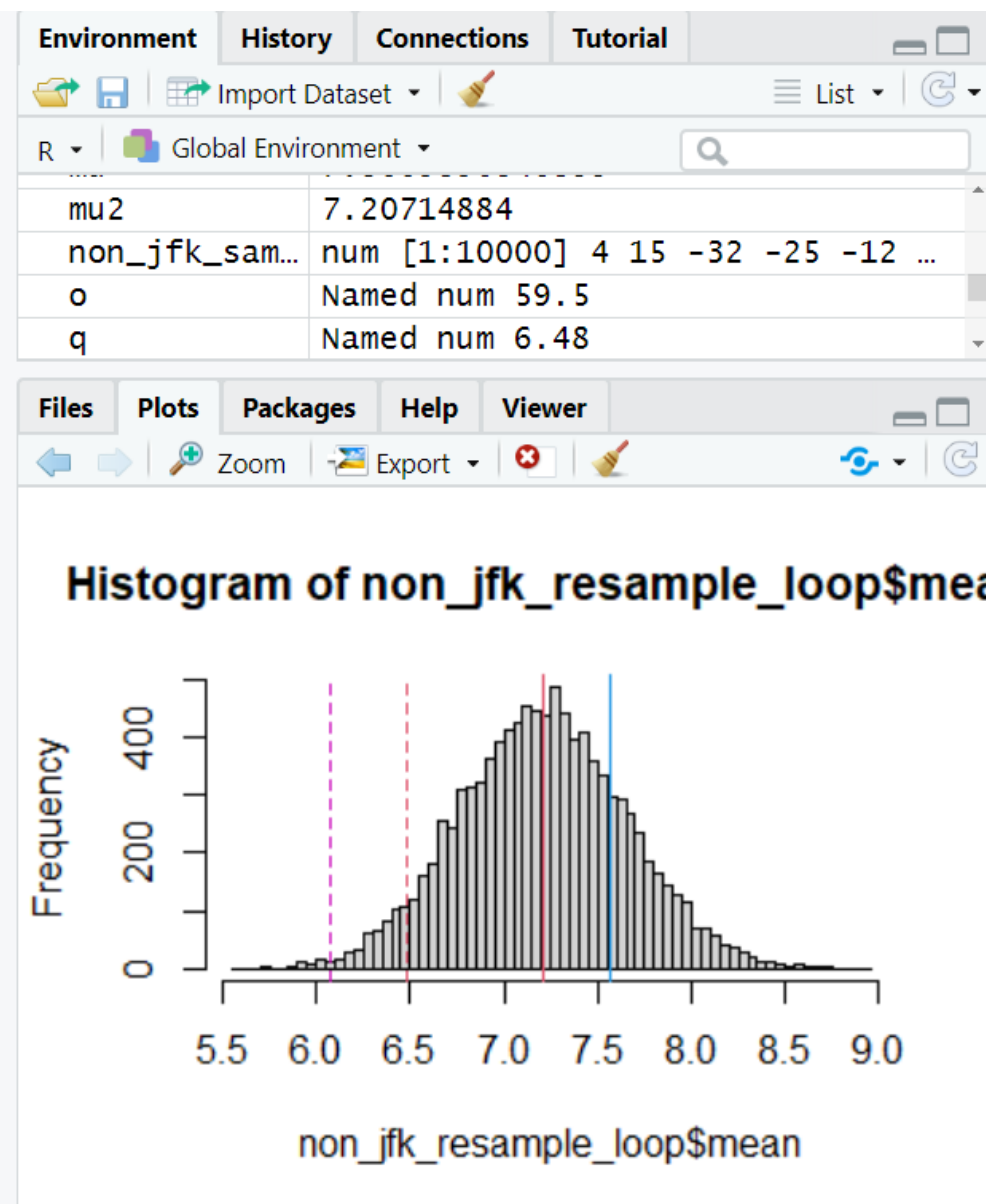
```
QM6c QM7 QM8.R x flights x Funktionen.R x jfk x
Source on Save Run Source
115 #aus diesem sample resampeln um die Breite des Konf-intervalls zu schÄ
116 set.seed(2021)
117
118 # resampeln mit der H_0: non-jfk
119 non_jfk_resample_loop <- do (10000)*mean(resample(non_jfk_sample)) #fo
120 mean(non_jfk_resample_loop$mean) #mean der means (grand mean)
121 sd(non_jfk_resample_loop$mean) #Standardfehler
122
123 hist(non_jfk_resample_loop$mean, breaks=50) #Histogramm
124 plot(density(non_jfk_resample_loop$mean,bw=1)) #GlÄttung Dichteplot
125
126 mu<-mean(non_jfk$arr_delay);mu # gesamte JFK
127 abline(v = mean(non_jfk$arr_delay), col = 4) #mean alle JFK
128
129 m_<-mean(non_jfk_resample_loop$mean);m_ # MW der Mittelwerte der samp1
130 abline(v = mean(non_jfk_resample_loop$mean), col = 2) #mean Stichprobe
131
```

```
Console Terminal x Jobs x
~/
> abline(v = mean(non_jfk$arr_delay), col = 4) #mean origins
> mu<-mean(non_jfk$arr_delay);mu # Mittelwert aller origins
[1] 7.566989
> hist(non_jfk_resample_loop$mean, breaks=50) #Histogramm
> m_<-mean(non_jfk_resample_loop$mean);m_ # MW der Mittelwerte der samples
[1] 7.207149
> abline(v = mean(non_jfk_resample_loop$mean), col = 2) #mean StichprobenMW
```



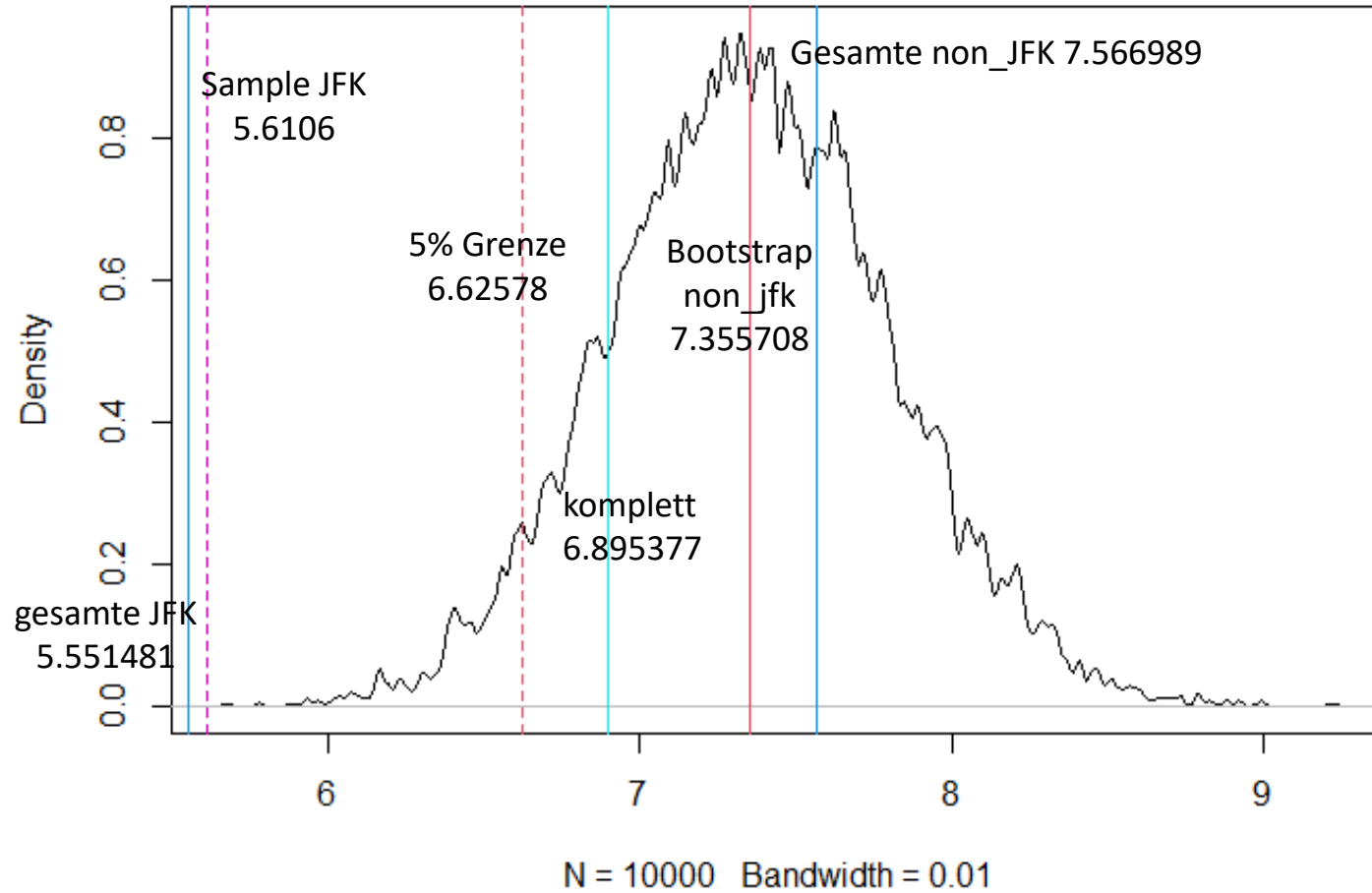
```
QM6c QM7 QM8.R* x flights x Funktionen.R x jfk x
Source on Save Run Source
131
132 ## 5% quantil (von links)
133
134 #H1: Die Flüge am JFK haben weniger Verspätung (Ankunft)
135 #H0: Die Flüge am JFK haben nicht weniger Verspätung (Ankunft)
136
137 q<-quantile(non_jfk_resample_loop$mean, prob=c(0.05));q
138 abline(v = q, col = 2, lty = 2)
139
140 abline(v = x_, col = 6, lty = 2)
141
142 x_ < q #Test: ist der mean jfk kleiner 5% der non_jfk?
143 paste("Der Stichprobenmittelwert von JFK beträgt",x_)
144 paste("Die 5% Grenze bei non_jfk liegt bei",q)
145
146 #####
146:1 # (Untitled) R Script
```

```
> x_ < q #Test: ist der mean jfk kleiner 5% der non_jfk?
5%
TRUE
> paste("Der Stichprobenmittelwert von JFK beträgt",x_)
[1] "Der Stichprobenmittelwert von JFK beträgt 6.0769"
> paste("Die 5% Grenze bei non_jfk liegt bei",q)
[1] "Die 5% Grenze bei non_jfk liegt bei 6.477565"
>
```





`density.default(x = non_jfk_resample_loop$mean, bw = 0.01)`



Sample JFK ist unterhalb der 5%- Grenze → Nullhypothese wird verworfen

**$H_1$ : Die Flüge am JFK haben weniger Verspätung (Ankunft)**

**$H_0$ : Die Flüge am JFK haben nicht weniger Verspätung (Ankunft)**

← gilt wenn  $< 5\%$

bleibt bei  $\geq 5\%$

