Motivation

- *Frage:* Zu welchem Signifikanzniveau hätten wir die Nullhypothese gerade noch verworfen?

  Antwort auf diese Frage liefert der p–Wert:

  Der ***p*–Wert** eines statistischen Tests gibt die Wahrscheinlichkeit an, unter der Nullhypothese einen extremeren Wert als die realisierte Teststatistik zu beobachten.

Definition

- Der p–Wert sagt aus, bis zu welchem vorgegebenen Signifikanzniveau $\mathcal{H}_0$ verworfen wird.

  Damit bewertet er, wie signifikant eine etwaige Abweichung von $\mathcal{H}_0$ ist.

  Entscheidungsregel: Verwerfe $\mathcal{H}_0$ falls der p–Wert kleiner als α ist.

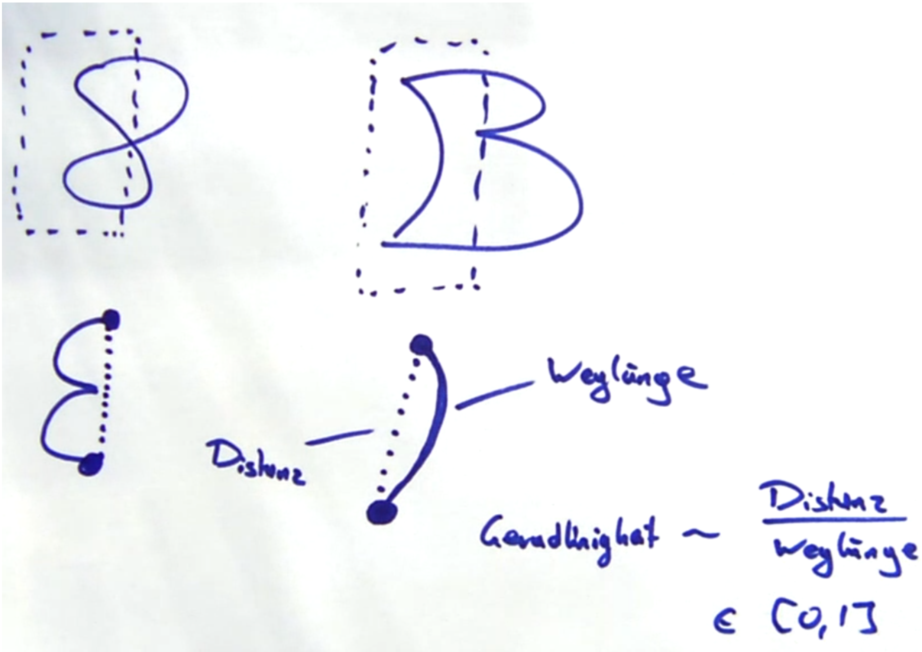
  Für eine statistisch saubere Argumentation, muss das gewünschte Signifikanzniveau α *vor* der Durchführung des Tests angegeben werden.

Misconceptions

- - **Warning!** A large p-value is not strong evidence in favor of H_0. A large p-value can occur because H_0 is true or H_0 is false but the test has low power.
  - **Warning!** p-value is not $P(H_0|Data)$, i.e. the p-value is not the probability that the null hypothesis is true
  - **Correct interpretation:** The p-value is the probability under H_0 of observing a value of test statistic the same as or more extreme than what we actually observed!

Beispiel

- Wir betrachten das Problem, die Zeichen '8' und 'B' voneinander zu unterscheiden.



- Es wird angenommen, dass die Geradlinigkeit der beiden Zeichen normalverteilt ist mit identer Standardabweichung $\sigma = 0.05$.
- Zur Beantwortung dieser Frage formulieren wir die Nullhypothese

$$\mathcal{H}_0: \quad \mu_X - \mu_Y = 0 \qquad \text{vs} \qquad \mathcal{H}_1: \quad \mu_X - \mu_Y \neq 0$$

und verwenden $\alpha = 0.05$.

Da es sich um ein Zweistichprobenproblem mit bekannter Varianz handelt, greifen wir auf einen Zweistichproben–Gauß–Test zurück.
-

```
> x <- c(0.86, 0.86, 0.93, 0.83, 0.83, 0.93, 0.81, 0.91,
         0.83, 0.88)
> y <- c(0.91, 0.89, 0.92, 0.87, 0.94, 0.99, 0.91, 0.96)
> n <- length(x)
> m <- length(y)
> sigma_x <- 0.05
> sigma_y <- 0.05
> z <- (mean(x)-mean(y))/sqrt(sigma_x^2/n + sigma_y^2/m)
> p.value <- 2 * pnorm(abs(z), lower.tail = FALSE)
> p.value
[1] 0.0167208
```

- Der p–Wert ist kleiner als $\alpha = 0.05$, also verwerfen wir die Nullhypothese

$$\mathcal{H}_0: \mu_X - \mu_Y = 0.$$

- Das Gerät ist also in der Lage, die beiden Zeichen voneinander zu unterscheiden.
-