<u>Einseitiger Hypothesentest</u>

Problematik: Von einem Bernoulli-Experiment (z.B. einer Produktion) kennt man $P(T) = p_0$.
Man vermutet/befürchtet, dass sich p_0 verändert hat. Dabei ist aber <u>nur die Richtung</u> der Veränderung bekannt, <u>nicht die Größe</u> der Veränderung (der veränderte Wert von P(T)).

Strategie: Man entnimmt eine Stichprobe der Länge n und <u>entscheidet aufgrund der Stichprobe</u>, ob die Vermutung (= **Gegenhypothese H_1**) stimmt oder auch nicht.
Haken an der Sache: Weil die Stichprobe ein Zufallsprozess ist, kann man nicht sicher sein, ob die getroffene Entscheidung zu 100% stimmt.

Linksseitiger Test: p_0 hat sich verkleinert

Eine Laplace-Münze zeigt Wappen mit einer Wahrscheinlichkeit von 0,5.
Die sog. **Nullhypothese H_0** lautet kurz: **$H_0 : p_0 = 0{,}5$**.
Bei einer besonderen Münze beobachtet man ein verringertes Auftreten von Wappen.
Man vermutet (= Gegenhypothese H_1), dass P(W) = p kleiner als erwartet ist, kurz: **$H_1 : p_1 < p_0$**
Eine Stichprobe von (willkürlich) 100 Würfen soll die Entscheidung bringen.
Wenn die **Anzahl von Wappen** bei diesen 100 Würfen (= **Testgröße T**) einen *vorher* willkürlich gewählten **kritischen Wert k**, z.B. k = 40 nicht überschreitet, ist man geneigt, die Gegenhypothese H_1 anzunehmen (und damit die Nullhypothese H_0 zu verwerfen).
Wenn die Stichprobe z.B. T = 40 oder T = 39 Wappen hat, entscheidet man sich für $H_1 : p_1 < 0{,}5$.
Zeigt die Stichprobe allerdings z.B. 48 Wappen (48 > k) zeigt, behält man H_0 bei.

Je kleiner man in diesem Beispiel den kritischen Wert k wählt, desto „schärfer" wird der Test. d.h. desto sicherer kann man eine stark verfälschte Münze als solche erkennen.
Allerdings gehen einem dabei leichte Abweichungen von p_0 durch die Lappen.
Eine Aufgabe des Hypothesentestes ist es unter anderem, dieses Dilemma in Zahlen zu fassen.

Wir gehen <u>immer</u> davon aus, dass sich p_0 nicht verändert hat, P(T) also bekannt ist.
(Nur so kann man überhaupt etwas berechnen.)
Das Ziel ist, einen **Zahlenwert für das Risiko einer Fehlentscheidung (gegen H_0)** zu ermitteln.
Dieses Risiko nennt man **Fehler 1. Art** oder **Risiko 1. Art**, bezeichnet wird es mit α.

Im Beispiel kann auch bei 100 Würfen einer idealen L-Münze z.B. genau 40mal Wappen auftreten.
Das geschieht mit der Wahrscheinlichkeit B(100; 0,5; 40) = 0,01084.
Entscheiden wir uns *bei diesem Ausgang* der Stichprobe gegen H_0, ist das Risiko 1. Art $\alpha = 0{,}01084$.
Nun kann man aber den Ausgang der Stichprobe im Voraus nicht vorhersagen.
Wir werden uns auch bei weniger Wappen gegen H_0 entscheiden und damit falsch entscheiden.
Insgesamt bergen *alle Wappenzahlen $\leq$ kritischer Wert k* das Risiko einer Fehlentscheidung.
Das Gesamtrisiko α' für eine Fehlentscheidung setzt sich aus all diesen Teilrisiken zusammen:

$$\boxed{\alpha = F(n; p; k)}$$

In unserem Beispiel also: = F(100; 0,5; 40) = 0,02844.
Konkret bedeutet dies:
In unserem Test mit 100 Würfen würden wir bei oftmaliger Durchführung in 2,8% der Fälle eine ideale L-Münze als verfälschten Münze einschätzen und damit einen Fehler erster Art begehen.
Wenn die untersuchte Münze allerdings verfälscht ist, könnten wir sie auch als ideal einschätzen.
Die Wahrscheinlichkeit, diese Fehlentscheidung zu treffen, nennt man **Fehler** oder **Risiko 2. Art β** .
Sie besteht darin, H_0 anzunehmen, obwohl H_0 nicht zutrifft.
Über die absolute Größe von β kann man keine Aussage treffen, weil der Wert von p_1 nicht bekannt ist.

Beide Fehlerarten hängen von der Wahl des kritischen Wertes k ab:
α nimmt im Beispiel ab, wenn k kleiner gewählt wird. Das lässt sich auch berechnen.
β nimmt im Beispiel zu, wenn k kleiner gewählt wird. Die Größe von β kann man nicht berechnen.