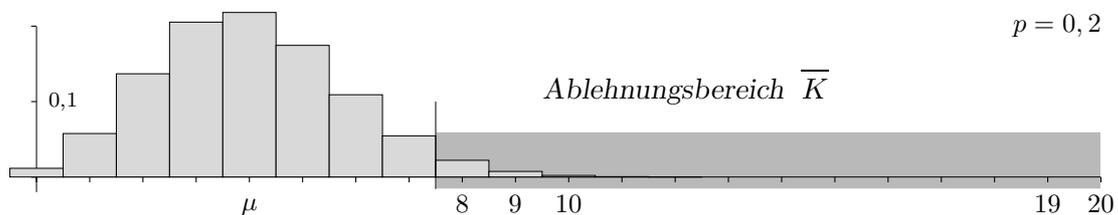


# Gütefunktion

Wir bezweifeln eine angeblich vorliegende (größtmögliche) Wahrscheinlichkeit von  $p = 20\%$  (z. B. verminderte Qualität in einer Lieferung, Wähler einer bestimmten Partei) und vermuten eine wesentlich höhere. Unser Interesse ist diese Vermutung zu belegen. Hierbei ist indirekt vorzugehen, d.h. es ist ein Bereich zu finden, der für die Wahrscheinlichkeit von  $p = 0,2$  gänzlich unwahrscheinlich ist. Wenn dann das Stichprobenergebnis in diesen Bereich fällt, können wir mit gutem Gewissen  $p = 0,2$  und dann auch  $p \leq 0,2$  als widerlegt ansehen. Natürlich können wir uns hierbei irren. Wir wählen den Stichprobenumfang  $n = 20$  und erlauben uns ein Irren in 5% der Fälle.

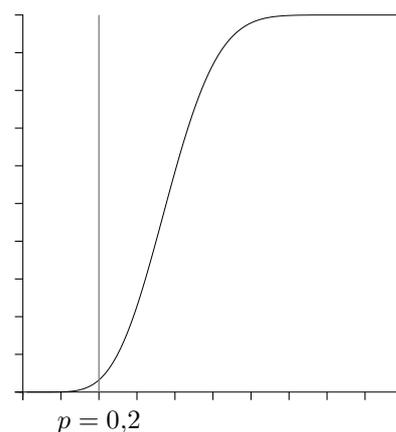
Sprechweisen:

Mit einem Signifikanztest soll die Nullhypothese  $H_0 : p = 0,2$  (genauer:  $H_0 : p \leq 0,2$ ) auf dem 5%-Niveau getestet werden. Es wird eine Entscheidungsregel für das Signifikanzniveau 5% gesucht. Die Gegenhypothese lautet  $H_1 : p > 0,2$ .



Der Fehler 1. Art, d. h. die Irrtumswahrscheinlichkeit (die Hypothese  $p = 0,2$  abzulehnen, obwohl sie richtig ist) beträgt  $\alpha = 3,21\%$ .

Um die Güte des Tests zu beurteilen, untersuchen wir für verschiedene angenommene  $p$ , mit welcher Wahrscheinlichkeit das Testergebnis in den kritischen Bereich  $\bar{K}$  fällt.



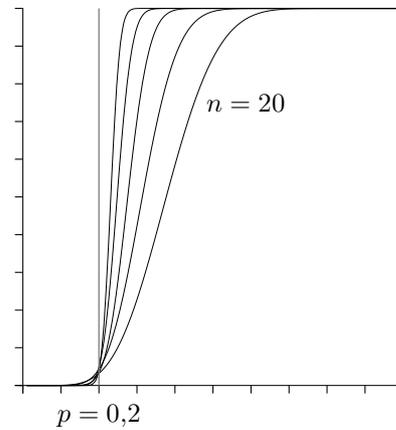
Die Funktion  $g(p) = P(\bar{K})$  heißt Gütefunktion.

Interpretieren Sie ihren Verlauf.

Wie können die Fehler 1. und 2. Art der Grafik entnommen werden?

# Gütefunktion

Nun variieren wir den Stichprobenumfang  $n$ . Der kritische Bereich ist von  $n$  abhängig. Der Test bleibt auf dem 5%-Niveau.



Die Gütefunktionen sind für  $n = 20, 40, 100, 500$  gezeichnet.

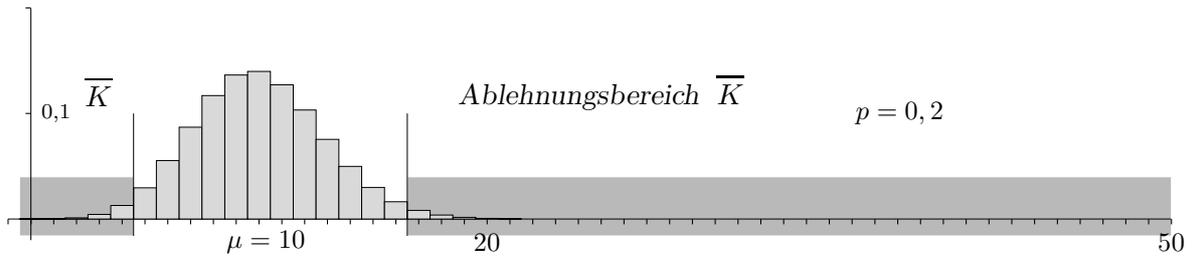
Interpretieren Sie die Grafik.

Gib es einen besten Test?

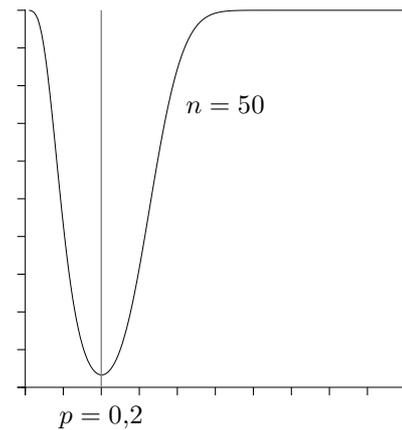
Wie verläuft die Gütefunktion eines zweiseitigen Tests?

# Gütefunktion eines zweiseitigen Tests

Mit einem Signifikanztest soll die Nullhypothese  $H_0: p = 0,2$  auf dem 5%-Niveau getestet werden. Die Gegenhypothese lautet  $H_1: p \neq 0,2$ . Der Ablehnungs- oder Verwerfungsbereich  $\bar{K}$  besteht nun aus zwei Bereichen, in die die Testgröße  $X$  jeweils mit einer (maximalen) Wahrscheinlichkeit von 2,5% fällt.

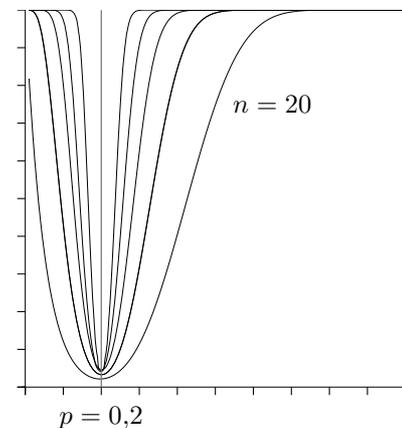


Interpretieren Sie die Grafik.



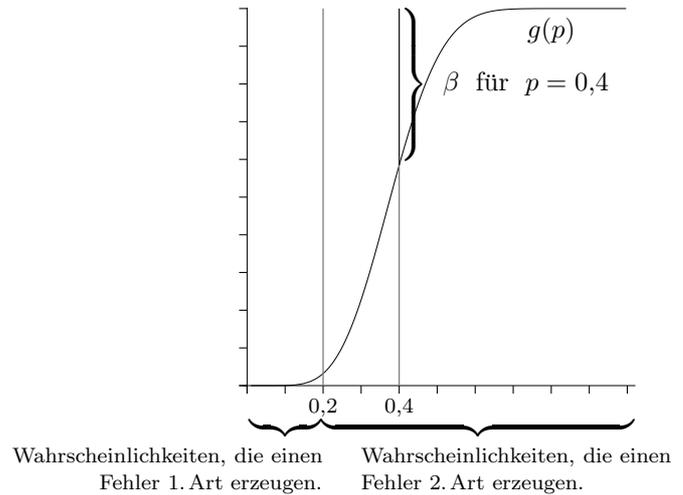
Die Gütefunktionen sind für  $n = 20, 50, 100, 200, 500$  gezeichnet.

Interpretieren Sie die Grafik.



## Gütefunktion Fehler 1. und 2. Art

Die Nullhypothese  $H_0: p \leq 0,2$  wird auf dem 5%-Niveau getestet,  $n = 20$ . Für  $X \geq 8$  wird die Nullhypothese verworfen.



Für  $p \leq 0,2$  ergeben sich die Fehler 1. Art direkt als Funktionswerte der Gütefunktion,  $\alpha = g(p)$ . Diese Fehler werden für  $p = 0,2$  maximal,  $\alpha = g(0,2) = 3,21\%$ .

Für  $p > 0,2$  ergeben sich die Fehler 2. Art als Differenz der Funktionswerte zu 1,  $\beta = 1 - g(p)$ .

Zusammengefasst (Notation etwas geändert):

Für einen Test sind die Nullhypothese  $H_0$ , der Ablehnungsbereich  $\bar{A}$  und der Nicht-Ablehnungsbereich  $A$  gegeben.

Die Gütefunktion des Tests

$$G(p) = P_p(\bar{A}) \quad 0 \leq p \leq 1$$

gibt die Wahrscheinlichkeit für das Ablehnen der Nullhypothese in Abhängigkeit von  $p$  an.

Die Operations-Charakteristik (OC) des Tests

$$O(p) = P_p(A) \quad (= 1 - G(p)) \quad 0 \leq p \leq 1$$

gibt den Fehler 2. Art in Abhängigkeit von  $p$  an.

Der OC-Graph entsteht aus dem Graphen von  $G(p)$  durch Spiegelung an der  $x$ -Achse und Verschiebung um 1 in  $y$ -Achsenrichtung.

Mit diesen Funktionen werden Fragen zum idealen Test, zur Verfälschtheit und zur Trennschärfe eines Tests untersucht.