KONFIDENZINTERVALLE

Einführung von Ac

<u>Problem:</u> („Schluss von der Stichprobe auf die Gesamtheit - Schätzen")
Von einer binomialverteilten Zufallsgröße X sei n (der Stichprobenumfang) gegeben, aber p (Erfolgswahrscheinlichkeit) unbekannt. Gesucht ist p bzw. ein Intervall von p-Werten.

<u>Konkreter formuliert:</u>
Eine Stichprobe liefere für die Zufallsgröße X den Wert x.
Welche Schlussfolgerung kann hieraus für p (Erfolgswahrscheinlichkeit) gezogen werden ?

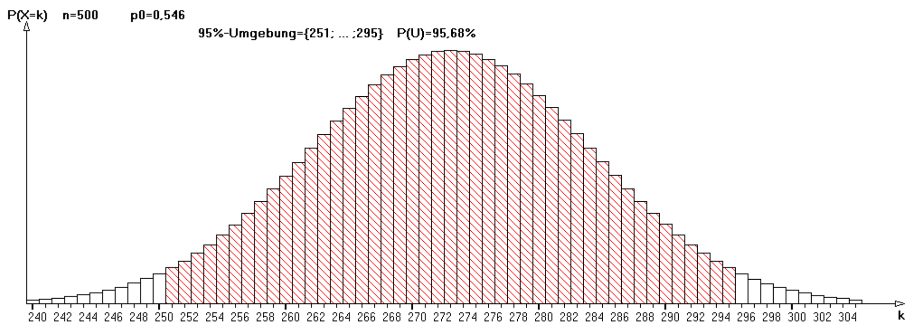
<u>**Beispiel:** (Wahlprognose)</u>

Man befragt 500 Personen, ob sie Kandidat A wählen werden. Dabei ergibt sich, dass 273 Personen voraussichtlich Kandidat A wählen . (Stichprobenergebnis also x = 273 und n = 500)
Schlussfolgerungen für die Erfolgswahrscheinlichkeit p des Kandidaten A:
p=0,546 ist ein guter Schätzwert für die Erfolgswahrscheinlichkeit. Warum ?

<u>**Frage 1:**</u> (wichtige Vorbetrachtung zur Erörterung des Themas)
Kann der Kandidat auf die Mehrheit der Stimmen hoffen ?

Beachte: Der Schätzwert garantiert noch keine Mehrheit, denn es ist unsicher, ob später wirklich 273 Personen den Kandidaten A wählen.

Wir betrachten dazu die Binomialverteilung für n=500; p=0,546 und schraffieren die 95%-Umgebung des Erwartungswertes μ=273, d.h. es bleiben links und rechts von μ jeweils 2,5% unschraffiert .
<u>Anmerkung:</u> Bei Meinungsumfragen ist es üblich, mit einer „**Sicherheitswahrscheinlichkeit**" von 95% zu arbeiten; daher ist die 95%-Umgebung des Erwartungswertes μ von Interesse .
(Anm.: Umgebungen von μ werden bei „Hypothesentests" auch als „**Annahmebereich A**" bezeichnet)



Die Grafik zeigt:

Gehen wir von einer Erfolgswahrscheinlichkeit p=0,546 aus, so wählen mit einer Sicherheitswahrscheinlichkeit von 95% zwischen 251 und 295 Personen den Kandidaten A.

Daraus kann man ersehen, dass Kandidat A mit 95%iger Sicherheitswahrscheinlichkeit mindestens 50% aller Stimmen bekommen wird. Aber Kandidat A kann sich der Mehrheit dennoch nicht ganz sicher sein, denn es bleibt ein Unsicherheitsfaktor von 5% .

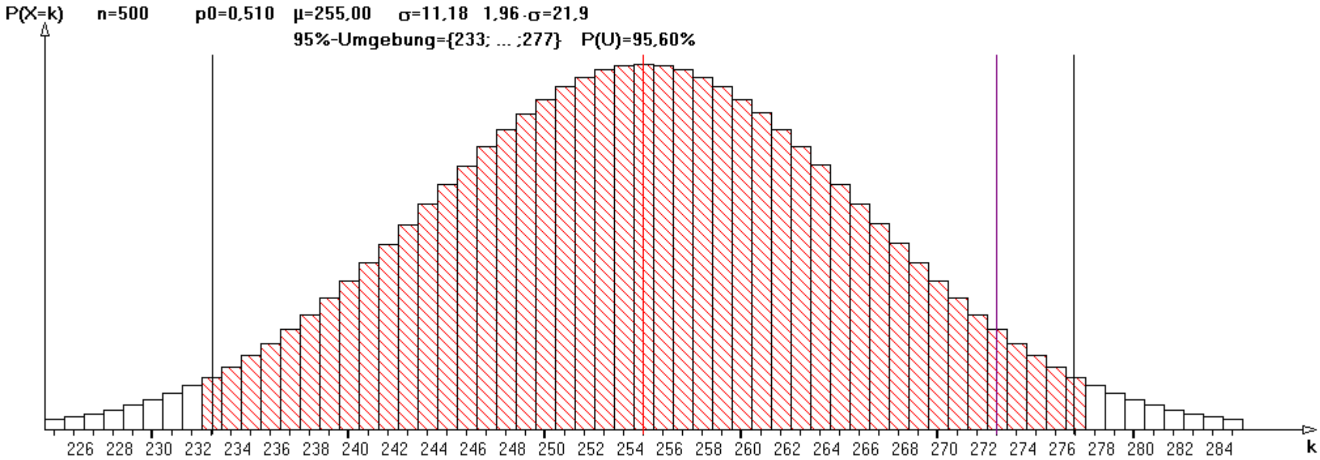
<u>**Frage 2:**</u> (<u>Dies ist die **eigentliche** Untersuchung zum Thema „Konfidenzintervalle"</u>)
Gibt es weitere p-Werte, in deren 95%-Umgebung das Stichprobenergebnis (hier: x=273) liegt ?

Natürlich gibt es ein ganzes Intervall von p-Werten, in deren Annahmebereich das Stichprobenergebnis x liegt. Man spricht hier von dem „**Vertrauensintervall (Konfidenzintervall)**" für die Erfolgswahrsch. p .

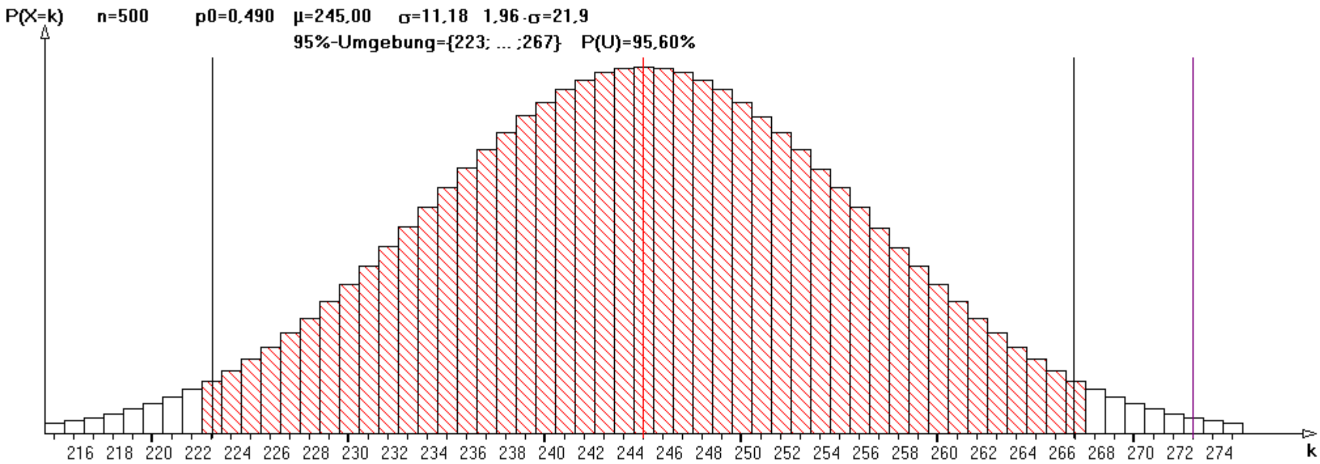
<u>**Aufgabe:**</u>

Bestimme zunächst die 95%-Umgebungen für p=0,51 bzw. p=0,49 .

Lösungen:



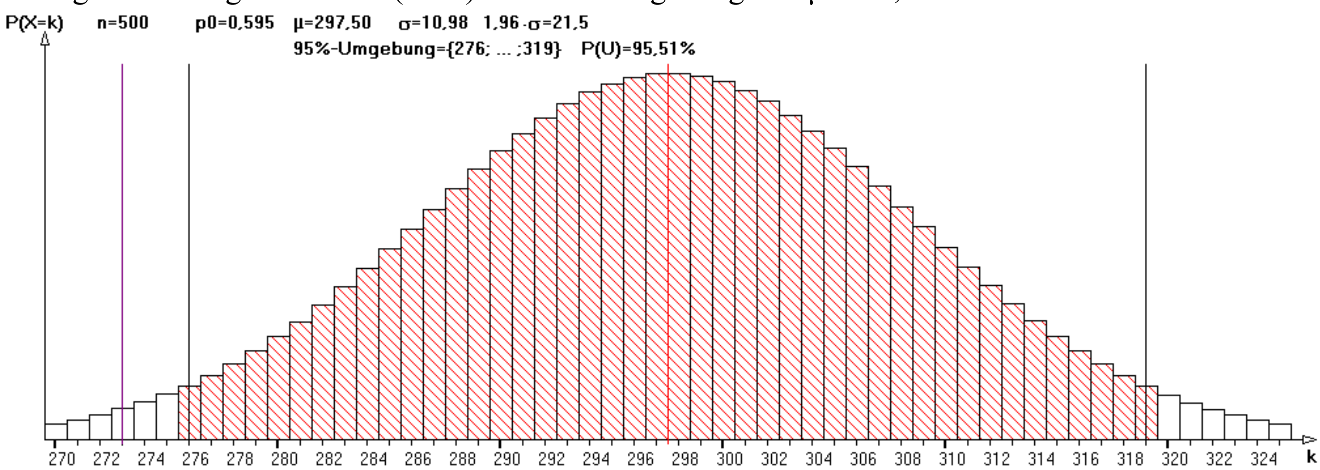
Hier liegt das Stichprobenergebnis 273 in der 95%-Umgebung von μ=255 .Daher ist p=0,51 **verträglich** mit dem Stichprobenergebnis 273. Verwendet man die 1,96σ-Umgebung und rundet zum Erwartungswert hin, so erhält man übrigens das Intervall {234; … ;276} mit einer Intervallwahrscheinlichkeit von 94,57% !



Hier liegt das Stichprobenergebnis 273 außerhalb der 95%-Umgebung von μ=245.
Daher ist p=0,49 **nicht verträglich** mit dem Stichprobenergebnis 273.

Zusatzaufgabe: Zeige, dass p=0,595 ebenfalls **nicht** verträglich mit dem Stichprobenergebnis 273 ist.

Lösung: 273 liegt außerhalb (links) der 95%-Umgebung von μ=297,5



Bezug zu Frage 2):

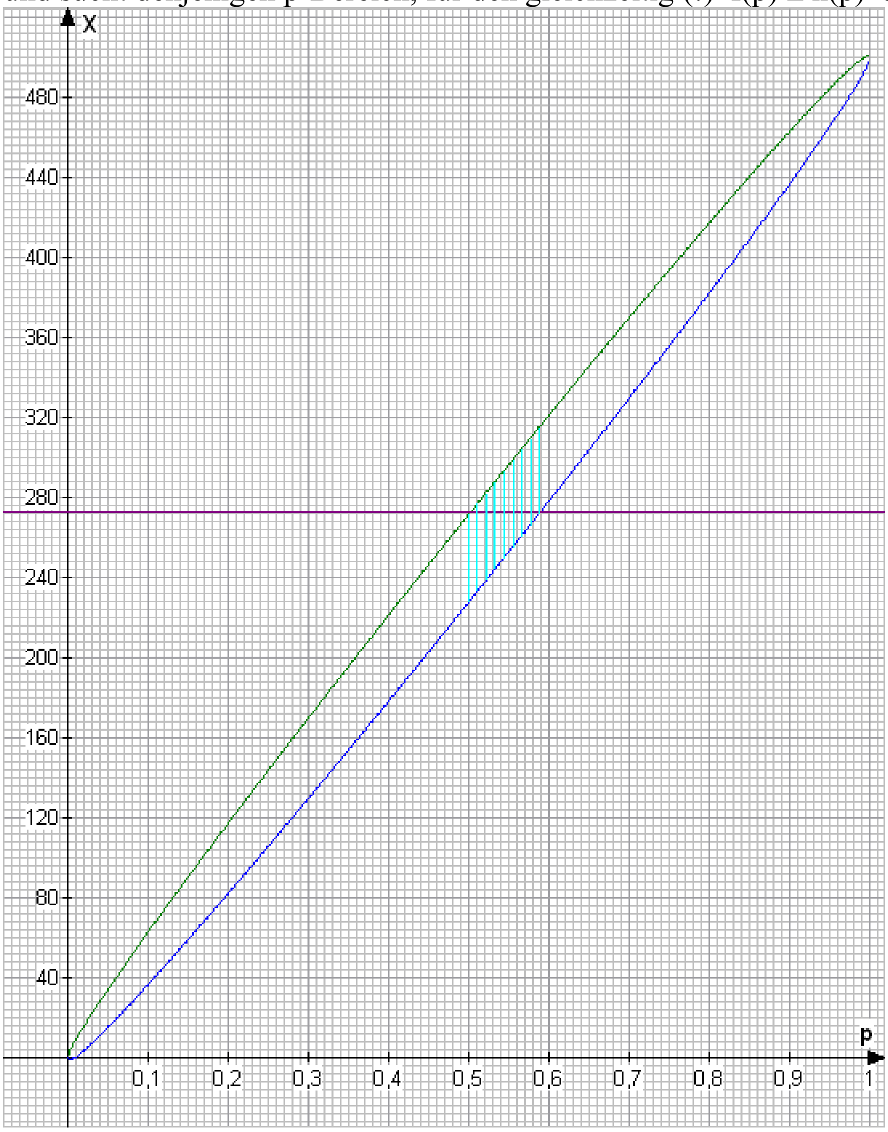
Gesucht ist ein **Intervall von p-Werten**, die verträglich mit dem Stichprobenergebnis x=273 sind, d.h. in deren jeweiliger Umgebung von μ das Stichprobenergebnis 273 liegt . Hierbei sind die jeweiligen Umgebungen (z.B. 95%) recht schwer zu bestimmen.

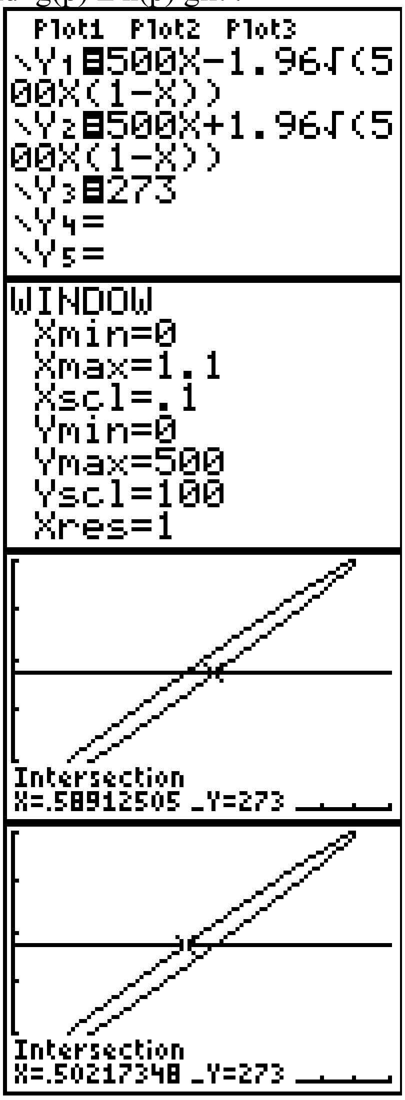
Leichter wird es mit einer Näherungsformel (σ-Regel), welche besagt:
Die 95%-Umgebung entspricht näherungsweise der 1,96σ-Umgebung von μ
Die p-Werte findet man demnach auf folgende Weise: $\mu - 1{,}96\sigma \leq x \leq \mu + 1{,}96\sigma$

Für unser Beispiel mit n=500 und x=273 ergibt sich also wegen μ=500p und $\sigma=\sqrt{500p(1-p)}$:

$$500p - 1{,}96\sqrt{500p(1-p)} \leq 273 \leq 500p + 1{,}96\sqrt{500p(1-p)}$$ ist also zu lösen !

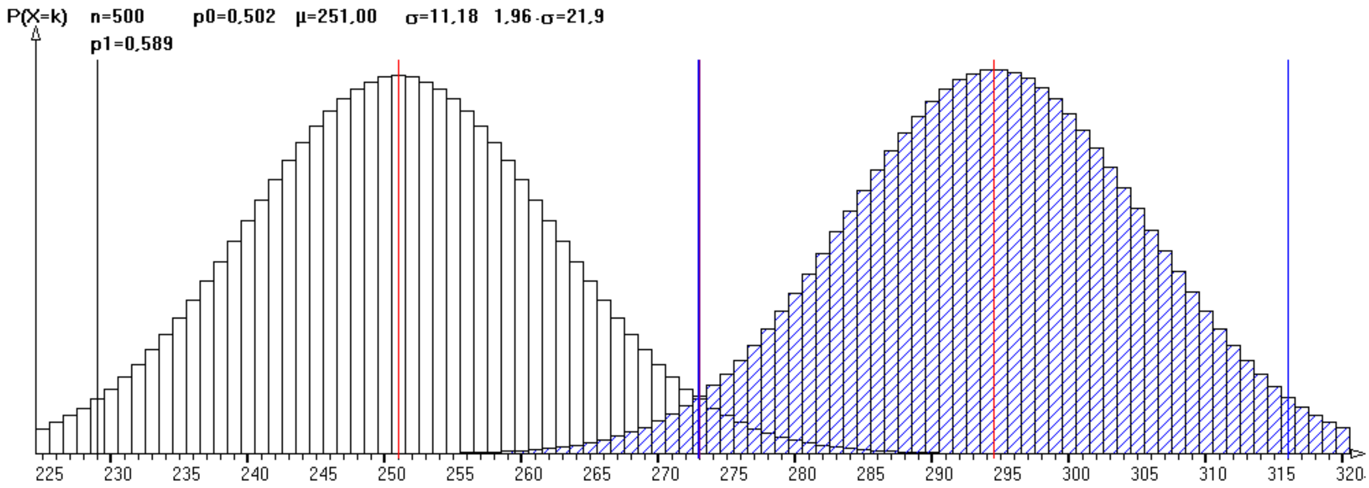
<u>Grafische Lösung für p:</u> Man lässt den GTR die drei Funktionen aus der Ungleichungskette zeichnen
$f(p) = 500p - 1{,}96\sqrt{500p(1-p)}$ $\quad$ $g(p) = 500p + 1{,}96\sqrt{500p(1-p)}$ $\quad$ $h(p) = 273$
und sucht denjenigen p-Bereich, für den gleichzeitig (!) $f(p) \leq h(p)$ und $g(p) \geq h(p)$ gilt .





Man erkennt, dass die Intervallgrenzen die Schnittstellen von f(p) bzw. g(p) mit h(p) sind.
Diese lassen sich beim GTR durch 2-malige Anwendung von **Intersect** (2nd CALC 5) bestimmen.
Also: Konfidenzintervall = [0,502 ; 0,589] .
<u>Ergebnis:</u> Alle p-Werte im Bereich [0,502 ; 0,589] sind verträglich mit dem Stichprobenergebnis 273.

Grafische Darstellung der Verteilungen der „Grenzwahrscheinlichkeiten“ p0=0,502 und p1=0,589:



Zu p0 gehört die weiße Verteilung und zu p1 gehört die blau schraffierte Verteilung .

Die 95%-Umgebung von p0=0,502 ist braun markiert (senkrechte Striche), während diejenige von p1=0,589 blau markiert ist. Die beiden rot markierten Stellen sind die Erwartungswerte der beiden Verteilungen.

An der Stelle k=273 fallen 3Markierungen zusammen, und zwar diejenigen des Stichprobenergebnisses x=273, die rechte Grenze der 95%-Umgebung von p0 und die linke Grenze der 95%-Umgebung von p1 !

Das Stichprobenergebnis liegt also auf dem gemeinsamen Schnittpunkt der 95%-Umgebungen von p0 und p1 und ist damit verträglich mit beiden Erfolgswahrscheinlichkeiten .

Algebraische Lösung (alternativ zur grafischen Lösung; eher für e-Kurse geeignet):

Ausgangspunkt zur Bestimmung der Intervallgrenzen sind die beiden Gleichungen:

$500p - 1{,}96\sqrt{500p(1-p)} = 273$ und $273 = 500p + 1{,}96\sqrt{500p(1-p)}$

Für die erste Gleichung folgt $500p - 273 = 1{,}96\sqrt{500p(1-p)}$

Für die zweite Gleichung entsprechend $273 - 500p = 1{,}96\sqrt{500p(1-p)}$

Es bietet sich jetzt an, die Gleichungen zu quadrieren, damit die Wurzel wegfällt.

$(500p - 273)^2 = (1{,}96\sqrt{500p(1-p)})^2$ sowie $(273 - 500p)^2 = (1{,}96\sqrt{500p(1-p)})^2$

$250000p^2 - 273000p + 74529 = 3{,}8416\cdot(500p\cdot(1-p))$ sowie

$74529 - 273000p + 250000p^2 = 3{,}8416\cdot(500p\cdot(1-p))$

Beide Gleichungen sind also identisch (!), daher formen wir die erste weiter um:

$250000p^2 - 273000p + 74529 = 1920{,}8\cdot(p-p^2) \Rightarrow$

$250000p^2 - 273000p + 74529 = 1920{,}8p - 1920{,}8p^2 \Rightarrow$

$251920{,}8p^2 - 274920{,}8p + 74529 = 0 \Rightarrow$

$p^2 - 1{,}0912985p + 0{,}29584298 = 0 \Rightarrow$

$p_1 \approx 0{,}50217$ $p_2 \approx 0{,}58913$

Ergebnis: (Rundung auf 3 Nachkommastellen)

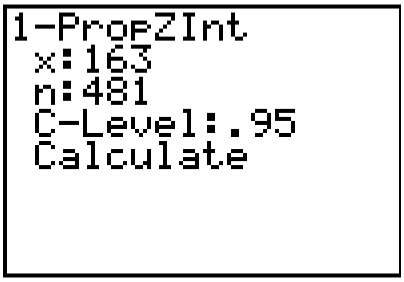
Das Konfidenzintervall ist also [0,502 ; 0,589] .

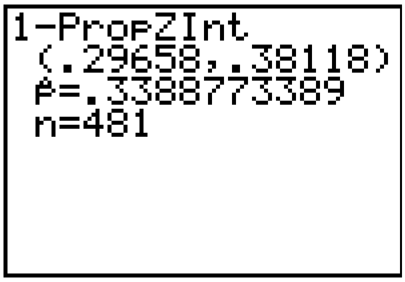
Verwendung der TI83-Funktion 1-PropZInt

Kommt es nicht auf die algebraische Lösung an, so kann man eine spezielle Funktion des GTR nutzen, um Konfidenzintervalle zu bestimmen:

Man wählt im **STAT TESTS** – Menü die **1-PropZInt** - Option .
Einzugeben sind x (Stichprobenergebnis), n (Umfang der Stichprobe) und C-Level (Sicherheitswahrscheinlichkeit) .
Ausgegeben wird dann (nach Anwahl von Calculate) das Konfidenzintervall.

Beispiel von oben: x=163 n=481





Das Ergebnis ist hier das Konfidenzintervall [0,297 ; 0,381] .

Dies weicht geringfügig von dem oben ermittelten Intervall ab (siehe Anmerkung unten) .

<u>Anmerkung:</u>
Eine genauere Untersuchung zeigt, dass das vom TI84 ermittelte Ergebnis **<u>falsch</u>** ist, denn es handelt sich hier offensichtlich nicht um eine Abweichung, die durch Rundung erzeugt wird, sondern man kann durch rechnerische Prüfung nachweisen, dass 1-PropZInt lediglich p = x/n <u>annimmt</u> (was nicht zulässig ist) und dann um diesen Wert herum die entsprechende Sigma-Umgebung berechnet !

Aufgaben und Beispiele:

1) Bestimme das 95%-Konfidenzintervall für eine Stichprobe vom Umfang 2000 mit dem Ergebnis 80.

2) Bei einer Untersuchung von 920 Personen hatten 402 die Blutgruppe A . Bestimme das 95%-Konfidenzintervall. Was sagt das Ergebnis aus ?

3) Beim Würfeln mit einem Quader erhielt man bei 481 Würfen 163-mal das Ergebnis 4. Bestimme das 95%-Konfidenzintervall.

4) Beim Würfeln mit einem Hexaeder erhielt man bei 3000 Würfen 468-mal das Ergebnis 6. Bestimme das 95%-Konfidenzintervall. Ist der Würfel gezinkt ?

Lösungen:

Zu 1): n=2000 und x=80. Es ergibt sich das Intervall [0,031 ; 0,049] .

Zu 2): n=920 und x=402. Es ergibt sich das Intervall [0,405 ; 0,469] .
Dies bedeutet, dass in der Bevölkerung (mit 95%iger Sicherheitswahrscheinlichkeit) zwischen 40,5% und 46,9% die Blutgruppe A besitzen .

Zu 3) Geometrische Lösung (ausführlich):
Gesucht ist das 1,96σ-Intervall $\quad np - 1{,}96\sqrt{np(1-p)} \le x \le np + 1{,}96\sqrt{np(1-p)}$
Teilt man die Ungleichungskette durch n, so rechnet man mit rel. Häufigkeiten h und erhält:

$$p - 1{,}96\sqrt{\frac{p(1-p)}{n}} \le \frac{x}{n} \le p + 1{,}96\sqrt{\frac{p(1-p)}{n}}$$

Stellt man h=X/n in Abhängigkeit von p dar, so ergibt sich ein „Ellipsendiagramm“ :

Die waagerechte Gerade ist die zur relativen Häufigkeit $h=x/n=163/481 \approx 0{,}3389$ gehörige .

Schneidet man die Ellipse mit dieser Geraden, so erhält man die Intervallgrenzen:
Man erhält für das Beispiel: [0,298 ; 0,382]

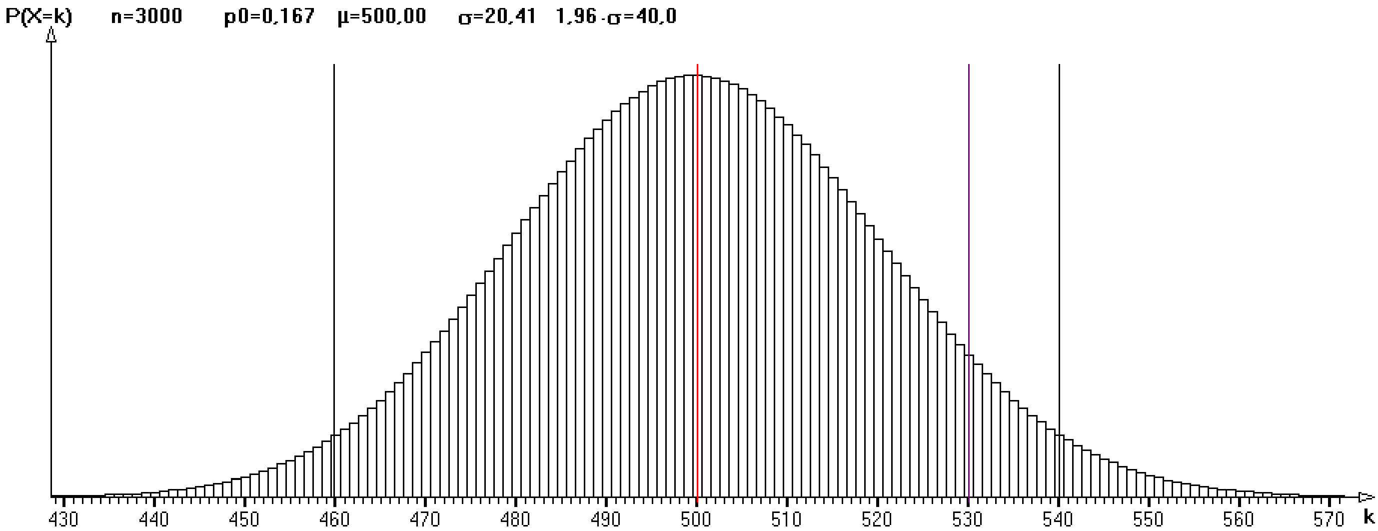
Zu 4) n=3000 und x=468. Es ergibt sich das Intervall [0,143 ; 0,169] .
Da die Wahrscheinlichkeit für das Auftreten einer 6 bei ca. 0,167 liegt und dieses Ergebnis im Konfidenzintervall enthalten ist, gibt es keinen Hinweis darauf, dass der Würfel gezinkt ist !

Notwendiger Stichprobenumfang n

Es interessiert noch die Frage, **wie groß der Stichprobenumfang mindestens sein muss**, wenn man eine Erfolgswahrscheinlichkeit p auf d% (z.B. 2%) genau schätzen will. D.h. Man gibt vor, dass x/n (Stichprobenergebnis/Stichprobenumfang) maximal d% vom wirklichen Wahrscheinlichkeitsanteil in der Gesamtheit abweichen soll.

Dazu ein Beispiel:
Man will bei einem Würfel die Wahrscheinlichkeit p% für die Augenzahl „6“ mit einer Genauigkeit von 2% schätzen. Das heißt, dass das Stichprobenergebnis x um höchstens 2% (von n) vom Erwartungswert μ für das Auftreten der „6“ abweichen darf. Die Sicherheitswahrscheinlichkeit für die Schätzung betrage 95%. <u>Gesucht</u> ist das für die Abweichung von 2% notwendige n (Stichprobenumfang) .

Damit man einen Eindruck von den Abweichungen bekommt, sind in der folgenden Grafik konkrete Werte für n (3000) und x (530) vorgegeben:



Rechnet man nach, so gilt hier: $x - \mu = 530 - 500 = 30$ und dies ist $\le 40 = 1{,}96\sigma \le 60 = 2\%$ von 3000.
Demnach ist in diesem Beispiel das n so groß gewählt, dass $x - \mu \le 2\%$ von n ist.

Losgelöst von den Zahlenbeispielen zeigt die Grafik folgendes (vorausgesetzt ist immer, dass das Stichprobenergebnis x innerhalb der 95% - Umgebung bzw. 1,96σ – Umgebung von μ liegt):
Liegt das Stichprobenergebnis x rechts von μ, so gilt: $x - \mu \le 1{,}96\sigma$.
Liegt das Stichprobenergebnis x links von μ, so gilt: $|x - \mu| \le 1{,}96\sigma$.
Es genügt daher, die Abweichung x-μ bzw. $|x - \mu|$ durch 1,96σ zu ersetzen !

Als Ansatz für die Genauigkeit von d% erhält man demnach: $\mathbf{1{,}96\sigma \le d\% \cdot n}$

Setzt man jetzt die Formel für σ ein und löst nach n auf, so ergibt sich eine Ungleichung für n in Abhängigkeit von p:

$$1{,}96\sigma \le d\% \cdot n \quad \Rightarrow \quad 1{,}96\sqrt{np(1-p)} \le d\% \cdot n \quad \Rightarrow \quad \frac{1{,}96}{d\%}\sqrt{n} \cdot \sqrt{p(1-p)} \le n$$

$$\Rightarrow \quad \frac{1{,}96}{d/100}\sqrt{p(1-p)} \le \frac{n}{\sqrt{n}} \quad \Rightarrow \quad \frac{196}{d}\sqrt{p(1-p)} \le \sqrt{n} \quad \Rightarrow \quad \left(\frac{196}{d}\right)^2 \cdot p \cdot (1-p) \le n$$

Vertauscht man jetzt noch die Seiten, so ergibt sich bei einer <u>95% - Sicherheitswahrscheinlichkeit</u>

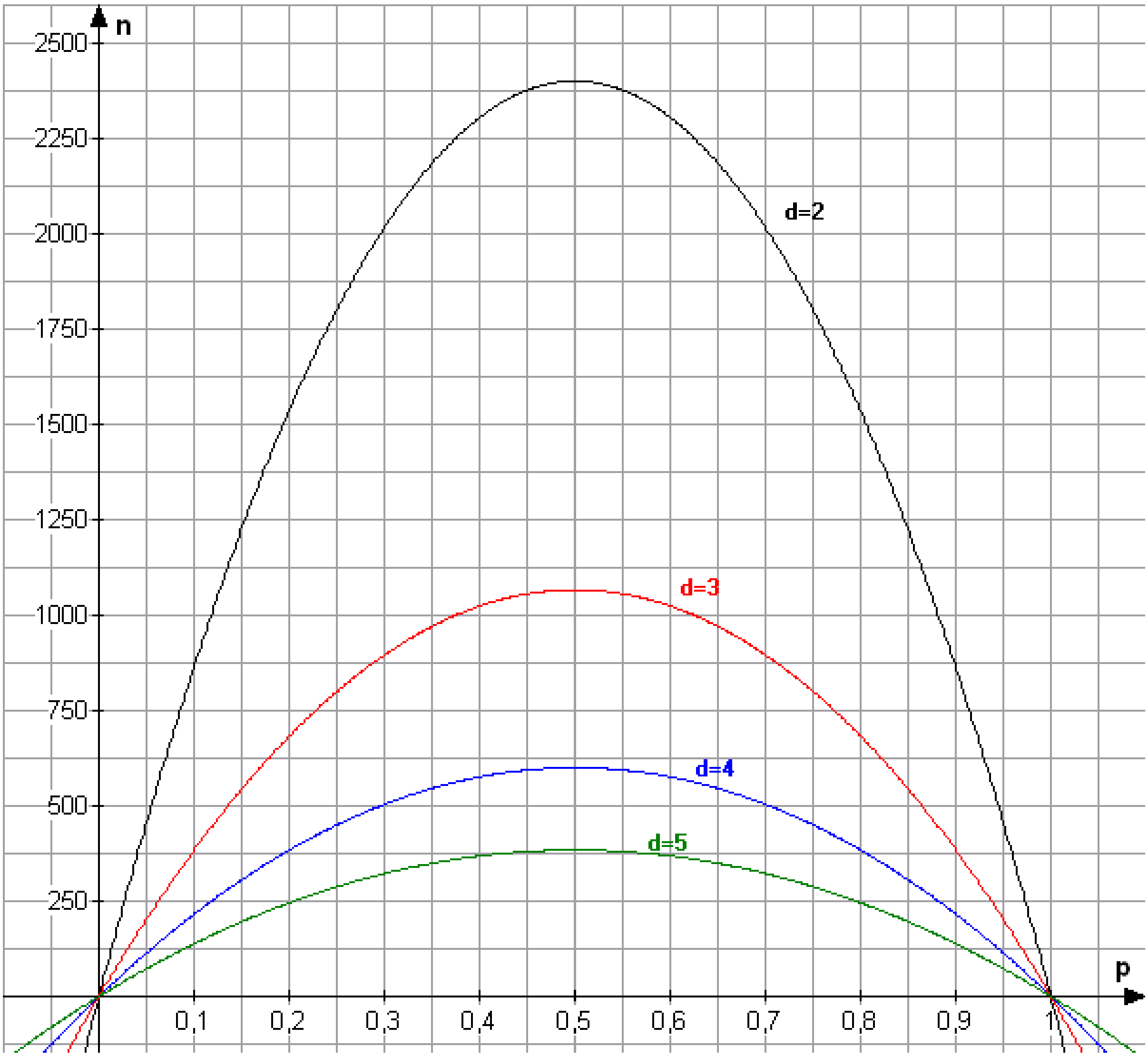
$$\boxed{n \ge \left(\frac{196}{d}\right)^2 \cdot p \cdot (1-p)} \qquad (*)$$

<u>Beispielrechnung</u> für eine Genauigkeit von 2% und $p=\frac{1}{6}$

$$n \geq \left(\frac{196}{2}\right)^2 \cdot \frac{1}{6} \cdot \left(1-\frac{1}{6}\right) \approx 1333{,}9$$

Also muss man mindestens einen Stichprobenumfang von **1334** wählen, damit man bei dem oben angegebenen Würfel die Wahrscheinlichkeit für die Augenzahl „6“ mit einer Genauigkeit von 2% schätzen kann.

Obige Formel (*) kann man auch grafisch (als Gleichung) darstellen:



Interpretation der Grafik ?

<u>Hinweis:</u>
Liegt keine Schätzung für die Erfolgswahrscheinlichkeit p vor (und dies ist meistens der Fall), so geht man von p=0,5 aus. Warum ?
Eine Schätzung könnte z.B. bei einer Wahlprognose vorliegen, wenn man aus Erfahrung weiß, dass eine Partei immer auf ca. 5% der Stimmen kommt .