3. Approximation der Binomialverteilung

durch die Normalverteilung

**Bsp1:**

Ein Spielwürfel wird 612 - Mal geworfen. Die Zufallsvariable beschreibt die Anzahl der Sechsen.

Die Zufallsvariable ist binomialverteilt mit und .

Die Zufallsvariable sei normalverteilt mit den Parametern und

In folgender Abbildung ist die Wahrscheinlichkeitsverteilung von als Histogramm und der Graph der Dichtefunktion der Zufallsvariable dargestellt.



Da die Balken des Histogramms die Breite 1 und die Höhe haben, ist ihr Flächeninhalt gleich der Wahrscheinlichkeit .

|  |  |
| --- | --- |
|  | Summe der Rechtecke des Histogramms der Binomialverteilung im Intervall |
|  | Fläche unter der Dichtefunktion der Normal­verteilung im Intervall |

Wir berechnen die Wahrscheinlichkeit :

**Bsp2:**

In einer Meinungsumfrage werden 400 Personen befragt, ob sie eine bestimmte Partei wählen werden. Le­diglich 50 Personen geben an, dass sie die Partei wählen werden.

Die Zufallsvariable , welche die Anzahl der Befragten wiedergibt, die angeben die Partei zu wählen, ist binomialverteilt mit und dem unbekannten Anteilswert .

|  |  |
| --- | --- |
| *Punktschätzung*: |  |

*Berechnung der unteren Grenze*  eines 95% - Konfidenzintervalls

Die Wahrscheinlichkeit, dass mindestens 50 Befragte angeben die Partei zu wählen, wird kleiner, wenn sich der Anteilswert verringert. Je kleiner der unbekannte Anteilswert ist, desto un­wahrscheinlicher ist es, dass mindestens 50 Befragte angeben die Partei zu wäh­len.

Somit ist die untere Grenze der kleinste Anteilswert, bei dem noch angenommen wird, dass mindestens 50 Befragte angeben die Partei zu wählen. Für kleinere Werte von er­scheint dies zu unwahrscheinlich, da die Wahrscheinlichkeit unter 2,5% fällt.

*Berechnung der oberen Grenze*  eines 95% - Konfidenzintervalls

Die Wahrscheinlichkeit, dass höchstens 50 Befragte angeben die Partei zu wählen, wird kleiner, wenn sich der Anteilswert erhöht. Je größer der Anteilswert ist, desto unwahr­scheinlicher ist es, dass höchstens 50 Befragte angeben die Partei zu wählen.

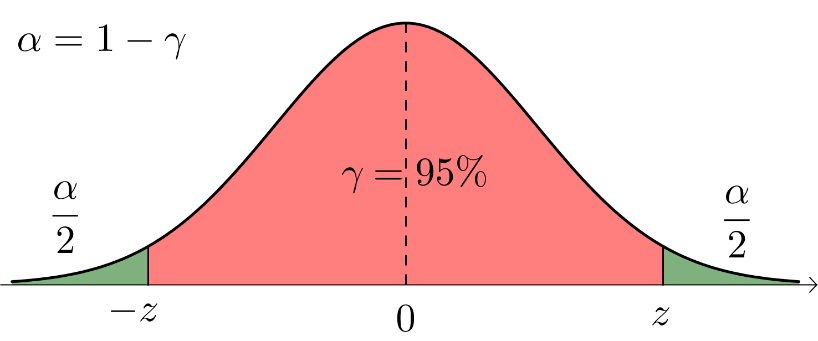
Somit ist die obere Grenze der größte Anteilswert, bei dem noch angenommen wird, dass höchstens 50 Befragte angeben die Partei zu wäh­len. Für größere Werte von er­scheint dies zu unwahrscheinlich, da die Wahrscheinlichkeit unter 2,5% fällt.

Im Intervall liegt der unbekannte Wähleranteil mit einer Wahr­scheinlichkeit von 95%.

*Approximation mit Hilfe der Normalverteilung:*

Wir nehmen die Zufallsvariable als normalverteilt mit und an.

Wir berechnen ein zum Erwartungswert symmetrisches Intervall, dass die Anzahl der Befragten, die angeben die Partei zu wählen. mit einer Wahrscheinlichkeit von enthält:



|  |
| --- |
| Näherungsformel zur Berechnung eine - Konfidenzintervalls |

Wenn diese Formel verwendet wird, sollte und sein. Trotzdem kann die Verwendung des Standard-Intervalls problematisch sein.

Für die gegebene Meinungsumfrage ist und

Die Näherungsformel ergibt das Konfidenzintervall , die exakte Berechnung mit der Binomialverteilung das Konfidenzintervall .

**Bsp3:**

In einer Meinungsumfrage unter 500 zufällig ausgewählten Personen zu Gentechnik in Lebensmitteln lehnen 80% der Befragten diese ab. Aufgrund dieses Ergebnisses gibt das Meinungsforschungsinstitut an, dass der Anteil der Personen, die Gentechnik in Lebens­mitteln ablehnen, zwischen 76% und 84% liegt.

Mit welcher Sicherheit kann diese Behauptung aufgestellt werden?

D.h. mit welcher Wahrscheinlichkeit liegt die Anzahl der Befragten die die Gentechnik ab­lehnen im vom Meinungsforschungsinstitut angegebenen Konfidenzintervall?

Berechnung der Überdeckungswahrscheinlichkeit mit der Binomialverteilung:

Mit einer Wahrscheinlichkeit von 97,8% kann gesagt werden, dass der Anteil der Personen, die Gentechnik in Lebens­mitteln ablehnen, zwischen 76% und 84% liegt.

Berechnung der Überdeckungswahrscheinlichkeit mit der Normalverteilung:

Mit einer Wahrscheinlichkeit von 97,5% kann gesagt werden, dass der Anteil der Personen, die Gentechnik in Lebens­mitteln ablehnen, zwischen 76% und 84% liegt.

**Bsp4:**

Berechne die Überdeckungswahrscheinlichkeiten für die in Bsp2 ermittelten Konfidenz­intervalle mit Hilfe der Binomialverteilung.

*Beachte:*

Die Berechnung mit Hilfe der Näherungsformel liefert das Konfidenzintervall mit einer Überdeckungswahrscheinlichkeit von 0,94178.

Die Überdeckungswahrscheinlichkeit liegt somit unter den geforderten 95%.

Die exakte Berechnung mit Hilfe der Binomialverteilung liefert das Konfidenzintervall mit einer Überdeckungswahrscheinlichkeit von 0,95754.

**Bsp5:**

Eine Fluggesellschaft weiß aus Erfahrung, dass bei Langstreckenflügen 10% der gebuch­ten Tickets nicht in Anspruch genommen werden. Sie setzt für diese Flüge Flugzeuge vom Typ Airbus A380-800 mit 526 Sitzplätzen ein.

Wie viele Plätze dürfen überbucht werden, damit das Risiko Fluggäste zurückweisen zu müssen höchstens 1% beträgt?

Berechnung mit der Binomialverteilung:

Soll das Risiko, dass Fluggäste zurückgewiesen werden müssen höchstens 1% betragen so dürfen nicht mehr als 41 Plätze überbucht werden.

Berechnung mit der Standardnormalverteilung:

Die Gleichung besitzt die Lösungen und .