

Information 1:

Für eine binomialverteilte Zufallsgröße X kann nach Bestimmung des Erwartungswertes μ und der Standardabweichung σ zu vorgegebenen Wahrscheinlichkeiten (meist 90%, 95%, 99%) die so genannte $z\sigma$ -Umgebung um μ bestimmt werden, innerhalb derer alle Erfolgsanzahlen zusammen die vorgegebene Wahrscheinlichkeit haben. Meist sind diese Werte $\mu - z\sigma$ und $\mu + z\sigma$ aber keine natürlichen Zahlen. Es muss also entschieden werden, welche Erfolgsanzahlen k_1 ($k_1 \approx \mu - z\sigma$) und k_2 ($k_2 \approx \mu + z\sigma$) zu verwenden sind.

Binomialverteilung: sigma-Regeln

$$P(\mu - 1,64\sigma \leq X \leq \mu + 1,64\sigma) \approx 90\%$$

$$P(\mu - 1,96\sigma \leq X \leq \mu + 1,96\sigma) \approx 95\%$$

$$P(\mu - 2,58\sigma \leq X \leq \mu + 2,58\sigma) \approx 99\%$$

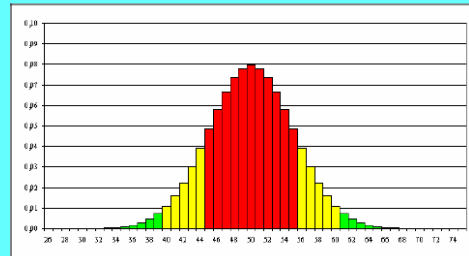


Abbildung aus der Präsentation von Strick: *Stochastik kompakt – worauf es ankommt*

Quelle: Folie 44 aus [<http://www.landrat-lucas.de/mint/stochastik/stochastik-ganz-kompakt.pdf>]

Wir werden dies an einem einfachen Würfel-Modell diskutieren:

Ein Würfel werde 600mal geworfen und das Auftreten der Augenzahl 6 (AZ 6) beobachtet.

Schluss von der Gesamtheit auf die Stichprobe:

Mit wie vielen Sechsern ist in 95% aller durchgeführten 600-er-Serien zu rechnen?

Die Berechnungen liefern (mit $n=600$; $p=1/6$; $z=1,96$; $\mu=100$; $V(X)=83,3$; $\sigma=9,13$; $z\sigma=17,98$)
 $\mu - z\sigma = 82,02$ und $\mu + z\sigma = 117,98$.

Nun sind die Grenzen k_1 und k_2 des Intervalls $[k_1 ; k_2]$ so zu bestimmen, dass sie sicher innerhalb des Intervalls $[\mu - z\sigma ; \mu + z\sigma]$ liegen, dieses Intervall ist also einzuengen. Es findet keine Rundung der beiden Werte statt, sondern es werden die geeigneten zugehörigen natürlichen Zahlen gewählt!

Im Beispiel sind also $k_1 = 83$ und $k_2 = 117$ zu wählen. Also ist $P(83 \leq X \leq 117) \approx 0,95$. Damit gilt: In 95% aller durchgeführten 600-er-Serien werden mindestens 83 und höchstens 117 Sechsern beobachtet.

Information 2:

Für Aussagen zur $z\sigma$ -Umgebung um μ (z. B.: „In 95% aller Versuche ...“) ist für das Intervall $[k_1 ; k_2]$ einzuengen (vgl. Buch Seite 444 – Runden zur sicheren Seite), um nur Erfolgsanzahlen zu benennen, die in der Summe wie gewünscht wahrscheinlich sind.

Zweiseitiger Hypothesentest:

Es wird vermutet, dass ein Würfel nicht regulär ist. Die Hypothese H_0 : *Der Würfel ist einwandfrei* wird getestet. Der 600er-Versuch erbringt 75 [115] Sechser. Wie ist zu entscheiden?

Die Berechnungen liefern wieder die oben gezeigten Werte.

Nun sind die Grenzen k_1 und k_2 des Intervalls $[k_1 ; k_2]$ so zu bestimmen, dass Erfolgsanzahlen außerhalb des Intervalls $[k_1 ; k_2]$ in der Summe wirklich unter 5% wahrscheinlich sind. Das Intervall $[\mu - z\sigma ; \mu + z\sigma]$ ist also aufzuweiten. Wieder findet keine Rundung der beiden Werte statt!

Im Beispiel sind also $k_1 = 82$ und $k_2 = 118$ zu wählen. Also ist $1 - P(84 \leq X \leq 118) \leq 0,05 = 5\%$. Das heißt: In höchstens 5% aller durchgeführten 600er-Serien werden weniger als 84 oder mehr als 118 Sechsern beobachtet. Die Entscheidungsregel lautet also: *Verwerfe die Hypothese, dass der Würfel regulär ist, falls weniger als 84 oder mehr als 118 Sechser in einer 600er-Serie beobachtet werden.* Tritt ein solches Versuchsergebnis ein, ist es ein seltenes Ereignis. Verwirft man deshalb die Hypothese H_0 und bestätigt also die Vermutung, so irrt man dabei **höchstens** in 5% aller Versuche mit 600 Würfeln.

Im Beispiel ist bei $X=75$ die Hypothese zu verwerfen, bei $X=115$ ist keine Entscheidung möglich.

Information 3:

Für Aussagen zum so genannten ‚Annahmereich‘ und zur Entscheidungsregel („Verwerfe die Hypothese ..., falls ...“) ist für das Intervall $[k_1 ; k_2]$ aufzuweiten, damit die Irrtumswahrscheinlichkeit sicher unter dem Vorgabewert bleibt (also wirklich höchstens 5% beträgt bei Verwendung der 95%-Umgebung).