Würfeln

Wir würfeln n mal. Betrachte folgende Szenarien:

- (a) Sie gewinnen CHF 10 falls es in mehr als 20% der Fälle einen "1" gibt.
- (b) Sie gewinnen CHF 10 falls es in mehr als 15% der Fälle einen "1" gibt.
- (c) Sie gewinnen CHF 10 falls es in 15%-20% der Fälle einen "1" gibt.

In jedem Szenario, was ist besser für Sie: n=60 oder n=600 mal würfeln?

Würfeln

Sei $X_i = 1$ falls sie einen "1" würfeln, und $X_i = 0$ sonst.
Wir sind interessiert an $\overline{X_n}$, also der Anteil "1" in n Würfen.

Es gibt:
$\mathrm{E}(\mathrm{X_i}) = P(X_i = 1) = \frac{1}{6} = 16\frac{2}{3}\%$.
Also $E(\overline{X_n}) = E(X_i) = 16\frac{2}{3}\%$ und $Var(\overline{X_n}) = \frac{Var(X_i)}{n} = \frac{\frac{1}{6}*\frac{5}{6}}{n} = \frac{5}{36n}$.

Die wichtigste Information:
Der Erwartungswert von $\overline{X_n}$ ist $16\frac{2}{3}\%$ und die Streuung nimmt ab wenn n wächst.

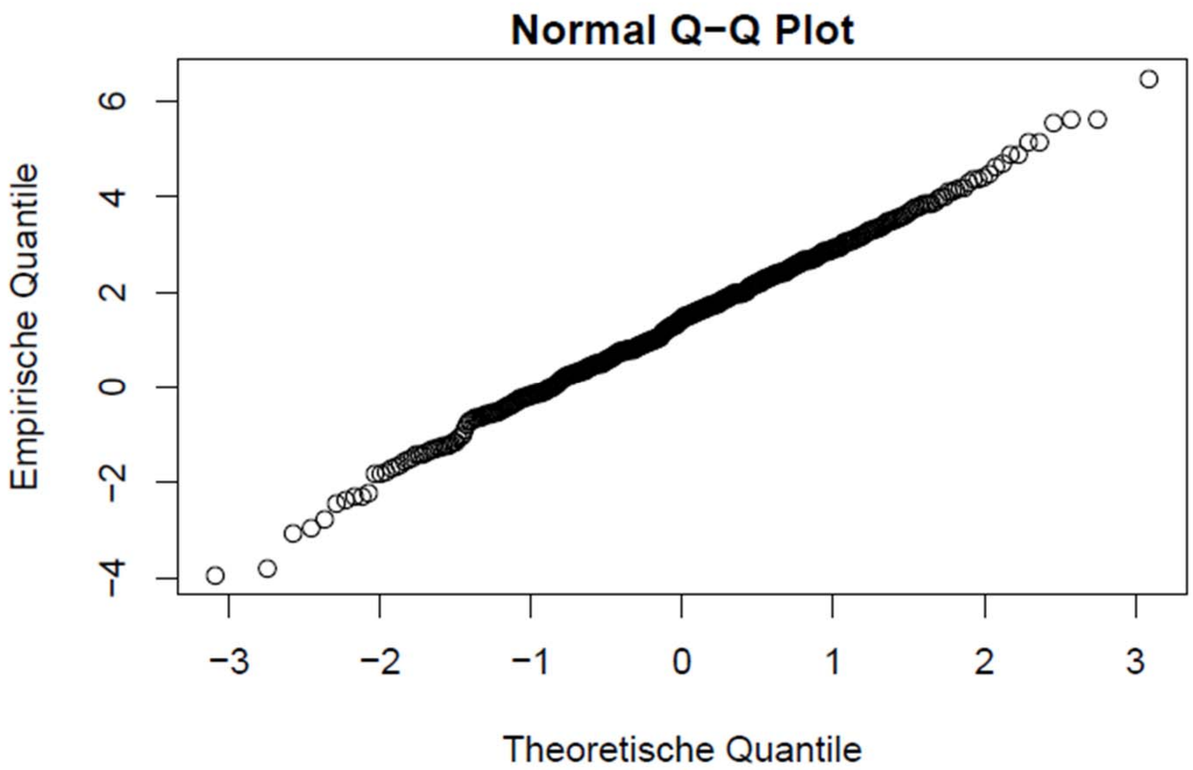
Würfeln

Wir betrachten jetzt die Szenarien:

(a) Sie gewinnen CHF 10 falls es in mehr als 20% der Fälle einen "1" gibt.
Hier gewinnen wir, wenn $\overline{X_n} > a$, wobei $a > E(\overline{X_n})$.
Viel Streuung hilft uns also, und wir wählen n so klein wie möglich, also $n = 60$.

(b) Sie gewinnen CHF 10 falls es in mehr als 15% der Fälle einen "1" gibt.
Hier gewinnen wir **nicht**, falls $\overline{X_n} < a$, wobei $a < E(\overline{X_n})$.
Viel Streuung schadet uns also, und wir wählen n gross, also $n = 600$.

(c) Sie gewinnen CHF 10 falls es in 15%-20% der Fälle einen "1" gibt.
Hier gewinnen wir **nicht**, falls $\overline{X_n} < a$ für $a < E(\overline{X_n})$ oder $\overline{X_n} > b$ für $\mathrm{b} > E(\overline{X_n})$.
Viel Streuung schadet uns also, und wir wählen n gross, also $n = 600$.



Betrachte die Aussagen

a) Die grösste Beobachtung im Datensatz ist ungefähr 3.
b) Der empirische Median ist kleiner als 0.

1. (a) Richtig / (b) Richtig
2. (a) Falsch / (b) Richtig
3. (a) Richtig / (b) Falsch
4. (a) Falsch / (b) Falsch
5. Keine Ahnung

QQ plot

- Die Beobachtungen im Datensatz sind genau die y-Werte der Punkte im QQ Plot. Die grösste Beobachtung ist also über 6 und Aussage a) stimmt nicht.

- Der QQ plot plottet jeweils (theoretisches q_α, empirisches q_α).
  Der theoretische Median der Standardnormalverteilung ist 0.
  Der y-Wert vom Punkt mit x-Wert 0 ist also der empirische Median.
  Dieser Wert liegt über 0 und Aussage b) ist deshalb falsch.