Herleitung des $\frac{1}{\sqrt{n}}$-Gesetzes

Für eine binomialverteilte Zufallsgröße X mit den Parametern n und p, dem Erwartungswert μ sowie der Standardabweichung σ gilt die Sigma-Regel für das 95 %-Prognoseintervall:

$$P(|\mu - X| \le 1{,}96\sigma) \approx 95\%.$$

Diese Näherung liefert brauchbare Ergebnisse, falls $\sigma > 3$ ist.

Anders formuliert bedeutet diese Regel: im Fall einer binomialverteilten Zufallsgröße mit $\sigma > 3$ liegen im Intervall $[\mu - 1{,}96\sigma;\ \mu + 1{,}96\sigma]$ nahezu 95 % der Zufallsausgänge.

Wir betrachten nun nicht die absoluten Häufigkeiten der binomialverteilten Zufallsgröße X, sondern als Zufallsgröße deren relative Häufigkeit $h_n = \frac{X}{n}$.

Da die Ereignisse $\{X_i = k\}$ und $\left\{\frac{X_i}{n} = \frac{k}{n}\right\}$ identisch sind, nimmt die Zufallgröße h_n ihre Werte $\frac{0}{n}; \frac{1}{n}; \ldots; \frac{n}{n}$ mit denselben Wahrscheinlichkeiten wie die Zufallgröße X an, d.h. h_n ist ebenfalls binomialverteilt mit den Parametern n und p.

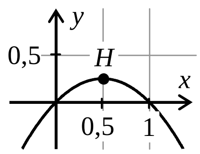
Für den Erwartungswert μ_h und die Streuung σ_h von h_n ergibt sich mithilfe der Regeln für die lineare Transformation:

- $\mu_h = E(h_n) = E\left(\frac{X}{n}\right) = E\left(\frac{1}{n} \cdot X\right) = \frac{1}{n} \cdot E(X) = \frac{1}{n} \cdot (n \cdot p) = p$
- $\sigma_h = \sqrt{\mathrm{Var}(h_n)} = \sqrt{\mathrm{Var}\left(\frac{X}{n}\right)} = \sqrt{\frac{1}{n^2} \cdot X} = \frac{1}{n} \cdot \sqrt{V\,\mathrm{ar}(X)} = \frac{1}{n} \cdot \sqrt{n \cdot p \cdot (1-p)}$
  $= \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} < \frac{0{,}5}{\sqrt{n}}$

Die Gültigkeit der verwendeten Abschätzung lässt sich wie folgt begründen:
Die Parabel zu der quadratischen Funktion mit der Gleichung

$$f(x) = -x^2 + x = x \cdot (1-x)$$

hat den Hochpunkt $H(0{,}5|0{,}25)$ (siehe Skizze).



Das bedeutet: Für alle x gilt $x \cdot (1-x) < 0{,}25 \iff \sqrt{x \cdot (1-x)} < \sqrt{0{,}25} = 0{,}5$.

Die Sigma-Regel für das 95 %-Prognoseintervall ist auch auf binomialverteilte Zufallsgröße h_n mit den Parametern n und p, dem Erwartungswert μ_h sowie der Standardabweichung σ_h anwendbar. Sie lautet:

$$P(|\mu_h - h_n| \le 1{,}96\sigma_h) \approx 95\% \iff P(|p - h_n| \le 1{,}96\sigma_h) \approx 95\%.$$

Dies bedeutet, dass mit nahezu 95 % Wahrscheinlichkeit gilt:

$$|p - h_n| \le 1{,}96\sigma_h < 1{,}96 \cdot \frac{0{,}5}{\sqrt{n}} = \frac{0{,}5}{\sqrt{n}} < \frac{1}{\sqrt{n}}.$$