Model für Meinungsumfragen
(z.B. Deutscher Einsatz in Afghanistan)

Jedes Mitglied der Bevölkerung könnte entweder dafür oder dagegen sein. Nehmen wir an, dass es N Personen insgesamt gibt, wovon eine Proportion p dafür sind und der Rest, $(1-p)$, alle dagegen sind.

Es wird eine zufällige Stichprobe der Größe n genommen. Sei X die Anzahl Personen aus n, die dafür sind.

Model 1 Hypergeometrisch

Model 2 Binomial

Model 3 Normal

Wie sehen die Resultate aus der Stichprobe aus, wenn die befragte Bevölkerung die Proportion $p_1 \neq p$ hat?

Wie soll das Modell erweitert werden, um Leute, die keine Angabe machen, zu berücksichtigen?

Bernoulli Experimente

Ein Experimenten mit Resultat von entweder 1 oder 0 wird als Bernoulli Experiment bezeichnet. Eine Reihe von n solchen unabhängigen Experimenten, wo für alle

$$P(X_i = 1) = p$$

ist ein Binomial Experiment mit

$$X = \sum_{i=1}^{n} X_i \sim B(n, p)$$

Eine Reihe von solchen unabhängigen Experimenten, die endet, wenn die Anzahl von Einsern r wird ist ein Negativbinomial Experiment mit

$$Y \sim NB(r, p)$$

wo Y die Anzahl der Bernoulli Experimente ist.

Gegeben, dass Sie jemanden beobachten, der mit dem 12. Versuch stoppt, wo er den 4. Einser bekommt, wie sind die Wahrscheinlichkeiten unter den zwei Modellen zu vergleichen?

Modelle für Ranking

z.B. Sportler, Schulen, Krankenhäuser werden in Ranking-listen aufgeführt.

Welche Eigenschaften sollten solche Rankings haben? Vergleichen wir Rang i mit Rang j für Sportler. Sei

$$p_{ij} = P(i \text{ schlägt } j \text{ im nächsten Spiel})$$

Wir könnten verlangen, dass

$$0 \leq p_{ij} \leq 1$$

$$p_{ij} > p_{ji} = 1 - p_{ij} \quad \text{für } i < j$$

$$p_{ij} < p_{ij+1}$$

$$p_{ij} > p_{i+1j}$$

$$p_{ij} > p_{i+1j+1}$$

Aus

$$\text{Odds}_{ij} = \frac{p_{ij}}{1 - p_{ij}} = \frac{r_j}{r_i}$$

bekommen wir

$$p_{ij} = \frac{r_j}{r_i + r_j}$$

$$r_i = i \Rightarrow p_{ij} = p_{2i2j}$$

Aus

$$p_{ij} \propto \frac{r_j}{r_i}$$

bekommen wir

$$p_{ij} = \frac{r_j^2}{r_i^2 + r_j^2}$$

$$r_i = i \Rightarrow p_{ij} = p_{2i2j}$$

Qualitätskontrolle

Sei μ das angestrebte Produktgewicht in einer Fabrik. Es werden Stichproben der Größe n jede Stunde i erhoben und das durchschnittliche Gewicht, $\bar{X}_{i,n}$ berechnet.

$$P(l_1 < \bar{X}_{i,n} < l_2) = ?$$

$$P(|\bar{X}_{i+1,n} - \bar{X}_{i,n}| > d) = ?$$

Was ist die Verteilung von

$$Q_{i,m} = \sum_{k=1}^{m} (\bar{X}_{i+k,n} - \mu) \quad ?$$

und welche Grenzen $\pm g_i$ sollte man dafür setzen, so dass

$$P(|Q_{i,m}| > g_i \mid |Q_{j,m}| < g_j \; \forall j < i) = \alpha_i$$