

Statistik I — bis jetzt

K 1 Einführung

K 2 Beschreibende Statistik

K 3 Graphiken

K 4 Schätzen

K 5 Testen

5.1 Tests im allgemeinen

5.2 Binomial Tests

5.3 Normal Test

5.4 Andere Mittelwert-Tests

und jetzt

5.5 Optimalität von statistischen Tests

5.5 Optimalität von statistischen Tests

Einfache und zusammengesetzte Hypothesen

Eine Hypothese wie $\theta = \theta_0$ ist eine einfache Hypothese, weil es die Verteilung vollständig bestimmt. Hypothesen wie $\theta \leq \theta_0$ werden zusammengesetzte Hypothesen genannt.

z.B.

$$H_0 : \theta = \theta_0 \qquad H_A : \theta = \theta_1$$

$$H_0 : \theta = \theta_0 \qquad H_A : \theta \neq \theta_0$$

$$H_0 : \theta \leq \theta_0 \qquad H_A : \theta > \theta_0$$

$$H_0 : \theta_1 \leq \theta \leq \theta_2 \qquad H_A : \theta < \theta_1 \text{ oder } \theta > \theta_2$$

Einseitige und zweiseitige Tests

Wenn die Hypothesen Alternativen in nur einer Richtung berücksichtigen (wie im dritten Beispiel), spricht man von einem einseitigen Test. Das führt dazu, dass ein einseitiger Test vom Niveau α einen anderen Ablehnungsbereich hat als ein zweiseitiger Test vom Niveau α . Vorsicht!

Beurteilungskriterien für statistische Tests

Testniveau α

Für einen Test des Niveaus α gilt $g(\theta) \leq \alpha \quad \forall \theta \in H_0$

Unverfälschtheit

Für einen unverfälschten Test des Niveaus α gilt

$$\forall \theta \in H_A$$

(Die Wahrscheinlichkeit H_A abzulehnen, wenn diese nicht vorliegt, sollte mindestens so groß sein, als wenn sie vorgelegen hätte.)

Konsistenz

Mit wachsender Stichprobengröße n soll der Test immer besser werden.

Optimalität

Sei $g(\theta), g_1(\theta)$ Gütefunktionen von Tests des Niveaus α , dann ist der Test mit Gütefunktion $g(\theta)$ der gleichmäßig beste Test vom Niveau α wenn

für $g_1(\theta)$ eines beliebigen anderen Tests vom Niveau α .

Optimale Tests

Neyman-Pearson

Neyman-Pearson Lemma für zwei einfache Hypothesen

$$H_0 : \quad \text{Dichte } f_0(\underline{x})$$

$$H_A : \quad \text{Dichte } f_A(\underline{x})$$

R sei irgendeine Region mit

$$P_0(\underline{x} \in R) \leq \alpha \quad \text{d.h.} \quad \int_R f_0(\underline{x}) d\underline{x} \leq \alpha$$

Nehmen wir an, es gebe eine Region

$$R^* = \{\underline{x} : f_A(\underline{x})/f_0(\underline{x}) \geq k\} \text{ mit } P_0(\underline{x} \in R) = \alpha$$

wo k so groß wie möglich ist, dann haben wir

$$P_A(\underline{x} \in R^*) \geq P_A(\underline{x} \in R)$$

Dieser Test (H_0 ablehnen, wenn $\underline{x} \in R^*$) ist der optimale Test des Niveaus α und R^* ist die sogenannte kritische Region.

Neyman-Pearson Beweis

Sei \bar{R}^* und \bar{R} die Komplemente von R^* bzw. R .

$$P_{\theta_A}(R^*) - P_{\theta_A}(R) = \int_{R^* \cap \bar{R}} f_A(x) dx - \int_{\bar{R}^* \cap R} f_A(x) dx$$

$$\text{Für } R^* \cap \bar{R} \quad f_A(x) \geq k f_0(x)$$

$$\int_{R^* \cap \bar{R}} f_A(x) dx \geq k \int_{R^* \cap \bar{R}} f_0(x) dx$$

Ähnlicherweise gilt

Deshalb haben wir, dass

$$P_{\theta_A}(R^*) - P_{\theta_A}(R) \geq k \left[\int_{R^* \cap \bar{R}} f_0(x) dx - \int_{\bar{R}^* \cap R} f_0(x) dx \right]$$

$$\begin{aligned} &= k \left[\int_{R^*} f_0(x) dx - \int_R f_0(x) dx \right] \\ &= k \left[P_{\theta_0}(R^*) - P_{\theta_0}(R) \right] \\ &\geq 0 \end{aligned}$$

Eine ähnliche Logik führt bei zusammengesetzten Hypothesen zum Likelihood-Quotienten-Test (Likelihood Ratio LR Test). Man entscheidet sich für die Alternative, wenn

einen Schwellenwert a überschreitet.

LR Test Beispiel

Betrachten wir die Hypothesen:

$$H_0 : X \sim N(\mu_0, \sigma_0^2)$$

$$H_1 : X \sim N(\mu_1, \sigma_1^2)$$

Wie sollte ein guter Test aussehen?

$$\bar{X} \sim N(\mu, \sigma_i^2/n)$$

$$R = \frac{f_1}{f_0} =$$

Für $\sigma_0^2 = \sigma_1^2 = \sigma^2$ und $\mu_1 > \mu_0$

$$R \geq a \Rightarrow \log R \geq \log a$$

$$\begin{aligned} \log R &= -\frac{n}{2\sigma^2} [(\bar{x} - \mu_1)^2 - (\bar{x} - \mu_0)^2] \\ &= -\frac{n}{2\sigma^2} [-2\bar{x}(\mu_1 - \mu_0) + (\mu_1^2 - \mu_0^2)] \\ &= \frac{n(\mu_1 - \mu_0)}{\sigma^2} \left[\bar{x} - \frac{(\mu_1 + \mu_0)}{2} \right] \end{aligned}$$

In diesem Fall hat der optimale Test eine erkennbare Form.

Wenn man die zwei Alternativen gleich behandelt, dann würde man nach Maximum Likelihood handeln und die entsprechende Hypothese auswählen.

Für $\sigma_0^2 \neq \sigma_1^2$

$$R \geq a \Rightarrow \log \frac{\sigma_1}{\sigma_0} R \geq \log \frac{\sigma_1}{\sigma_0} a$$

$$\Rightarrow -\frac{1}{\sigma_1^2}(\bar{x} - \mu_1)^2 + \frac{1}{\sigma_0^2}(\bar{x} - \mu_0)^2 \geq \frac{2 \log \frac{\sigma_1}{\sigma_0} a}{n}$$

\Rightarrow

Der Ablehnungsbereich eines optimalen Tests wird von der Lösung der quadratischen Gleichung abhängen.

KIs und Tests

Ist C ein Konfidenzbereich zum Niveau $1 - \alpha$, dann ist $\{\theta_0 \in C\}$ der Annahmebereich eines Tests von $H_0 : \theta = \theta_0$ gegen $H_1 : \theta \neq \theta_0$ zum Niveau α

Ist umgekehrt für jedes $\theta_0 \in \Theta$ ein (nichtrandomisierter) Test für $H_0 : \theta = \theta_0$ gegen $H_1 : \theta \neq \theta_0$ zum Niveau α gegeben, so läßt sich daraus ein Konfidenzbereich zum Niveau α gewinnen.

Bayes und Tests

Die ZV X hat die Dichte $f(x, \theta)$ und es muß eine apriori Verteilung für θ geben, $h(\theta)$. Die aposteriori Verteilung

$$h(\theta|x) \propto f(x, \theta)h(\theta)$$

enthält alle Informationen aus den Daten. Wir testen nicht. Statt dessen können wir Wahrscheinlichkeiten angeben:

$$P(\theta_1 < \theta < \theta_2) =$$