Neyman und Fisher

Original-Zitat von Neyman

(Kommentar von T. P. Speed zu: Samuels ML, Casella G, McCabe GP 1991 Interpreting blocks as random factors. JASA 86, 798-821.)

"I will now discuss the design of a field experiment involving plots ... In designing the experiment let us consider a field divided into m equal plots … To compare v varieties, we will consider that many sequences of numbers, each of them having two indices (one corresponding to the variety and one corresponding to the plot):

$$U_{i1}, U_{i2}, \ldots, U_{im} \ (i = 1, 2, \ldots v) \qquad (*)$$

[Neyman has already introduced a notion of true yield and U_{ik} is the true yield of variety i on plot k.]

The number

$$a_i = \frac{\sum_{k=1}^{m} U_{ik}}{m}$$

is the average of the numbers (*) and is the best estimate of the yield from the *i*-th variety on the field.

The goal of a field experiment which consists of the comparison of *v* varieties will be regarded as equivalent to the problem of comparing the numbers

$a_1, a_2, \ldots, a_v$

or their estimates."

Eine Identität

$X_{ij}(k)$ = "wahrer Ertrag" der k-ten Behandlung, der auf der j-ten Parzelle im i-ten Block erzielt werden kann

$$X_{ij}(k) = X_{\bullet\bullet}(k) + B_i(k) + \eta_{ij}(k)$$

wobei

$B_i(k) = X_{i\bullet}(k) - X_{\bullet\bullet}(k)$ = Effekt für die Fruchtbarkeit im i-ten Block,

$\eta_{ij}(k) = X_{ij}(k) - X_{i\bullet}(k)$ = Effekt für die Fruchtbarkeit der j-ten Parzelle im i-ten Block (Bodenfehler)

$\Rightarrow$ Identität

$$X_{ij}(k) = X_{\bullet\bullet}(k) + \left(X_{i\bullet}(k) - X_{\bullet\bullet}(k)\right) + \left(X_{ij}(k) - X_{i\bullet}(k)\right)$$

Gemessener Ertrag:

$$x_{ij}(k) = X_{ij}(k) + \varepsilon_{ij}(k)$$
$$= X_{\bullet\bullet}(k) + B_i(k) + \eta_{ij}(k) + \varepsilon_{ij}(k)$$

wobei

$\varepsilon_{ij}(k)$ = technischer Fehler

Zitat von Neyman in der engl. Übersetzung:

"Our purpose in the field experiment is to compare numbers such as $X_{\bullet\bullet}(k)$, or average true yields which our objects are able to give when applied to the whole field (Neyman 1935, p.111)."

Varianz-Kovarianz-Struktur im Modell mit unabhängigen Effekten (übliches Modell)

$$m_{ij} = \mu + b_j + f_{ij}$$

wobei

$b_j \sim N\left(0, \sigma_b^2\right)$ und
$f_{ij} \sim N\left(0, \sigma_f^2\right)$ ist.

$$m \sim MVN\left[\left(1_b \otimes 1_v\right)\mu, V\right] \quad ,$$

wobei *m* der nach Blöcken und Parzellen in Blöcken geordnete Vektor der m_{ij} ist und

$$V = I_b \otimes \left(vK_v\sigma_b^2 + I_v\sigma_f^2\right) \quad .$$

Man kann zeigen, dass dies äquivalent ist zu (John und Williams, 1995, p.136)

$$V = \xi_0 C_0 + \xi_1 C_1 + \xi_2 C_2 \quad , \qquad \text{(II)}$$

wobei

$$C_0 = K_b \otimes K_v \quad ,$$
$$C_1 = (I_b - K_b) \otimes K_v \quad ,$$
$$C_2 = I_b \otimes (I_v - K_v) \quad ,$$
$$\xi_0 = \xi_1 = v\sigma_b^2 + \sigma_f^2 \quad ,$$
$$\xi_2 = \sigma_f^2 \quad .$$

I_n = n-dimensionale Einheitsmatrix,
J_n = $n \times n$ Matrix mit Einsen,
$K_n = n^{-1} J_n$ und
m = der nach Blöcken und Parzellen in Blöcken geordnete Vektor der Werte m_{ij}.

Varianz-Kovarianz-Struktur im Randomisationsmodell (Neyman)

Das **Randomisationsmodell von Neyman** (ohne Messfehler) impliziert die Struktur

$$V = \xi_1 C_1 + \xi_2 C_2$$

Das **übliche Modell** impliziert die Varianz-Kovarianz-Struktur

$$V = \xi_0 C_0 + \xi_1 C_1 + \xi_2 C_2$$

Vergleich der beiden Modelle:

- Varianzen von Mittelwerten sind verschieden
- Varianzen einer Differenz aber immer gleich!

Fazit:

- **Blöcke immer als zufällig zu betrachten**, wenn Blöcke randomisiert wurden

- Randomisationsmodell und übliches Modell sind verschieden, aber sie liefern **dieselben Kontraste und Varianzen für Kontraste**

- Kann also immer mit üblichem Modell auswerten, sollte aber nur die Differenzen (Kontraste) und deren Standardfehler interpretieren

- **Standardfehler von Mittelwerten sollte man ignorieren**

- Der Hauptgrund, Blöcke zufällig zu nehmen, ist eine Nutzung der **Inter-Block-Information**. Das ist **nur lohnend, wenn man viele Blöcke hat**! Sonst besser, Blöcke fest, also auf die Inter-Block-Information verzichten.