Eins, zwei, drei: Stichprobenplanung leicht gemacht

Hans-Peter Piepho

Fachgebiet Biostatistik

Institut für Kulturpflanzenwissenschaften

Universität Hohenheim

One, two, three: Portable sample size in agricultural research

Hans-Peter Piepho[1,*], Doreen Gabriel[2], Jens Hartung[1], Andreas Büchse[3], Meike Grosse[4], Sabine Kurz[5], Friedrich Laidig[1], Volker Michel[6], Iain Proctor[3], Jan Erik Sedlmeier[7], Kathrin Toppel[8], Dörte Wittenburg[9]

1 Biostatistics Unit, Institute of Crop Science, University of Hohenheim, Stuttgart, Germany
2 Institute for Crop and Soil Science, Julius Kühn-Institut (JKI), Braunschweig, Germany
3 BASF SE, Ludwigshafen am Rhein, Germany
4 Research Institute of Organic Agriculture FiBL, Frick, Switzerland
5 Fachgebiet Pflanzenbau, Hochschule für Wirtschaft und Umwelt Nürtingen-Geislingen (HfWU), Nürtingen, Germany
6 Mecklenburg-Vorpommern Research Centre for Agriculture and Fisheries, Gülzow, Germany
7 Applied Entomology, Institute of Phytomedicine, University of Hohenheim, Stuttgart, Germany
8 Fachgebiet Tierhaltung und Produkte, Hochschule Osnabrück, Osnabrück, Germany
9 Research Institute for Farm Animal Biology (FBN), Institute of Genetics and Biometry, Dummerstorf, Germany

* corresponding author: piepho@uni-hohenheim.de

⇒ eingereicht bei *Journal of Agricultural Science*

Zweiter Teil in Arbeit:

Using a linear model package to compute precision and power and determine sample size

Zentrale Thesen

- Experimente sind komparativ
- Analyse kann immer mit paarweisen Vergleichen erfolgen
- Die elementare Größe für die Versuchsplanung ist der Standardfehler einer Differenz = Standard error of a difference (*SED*)
- Alle Versuchsplanung kann auf Basis des *SED* erfolgen

Die 1-2-3 Regel

- Spezifikation der angestrebten Genauigkeit kann direkt mit 1 $\times$ *SED* erfolgen

- Die Grenzdifferenz (LSD) ist etwa 2 $\times$ *SED /*

  Die erwartete halbe Breite eines Vertrauensintervalls ist etwa 2 $\times$ *SED /*

  Etwa 95% der geschätzten Differenzen weicht um höchstens 2 $\times$ *SED* von der wahren Differenz ab

- Die kleinste relevante Differenz, die mit einer Teststärke von 85% nachgewiesen werden kann, beträgt 3 $\times$ *SED*

Genauigkeitsvorgabe in Form des *SED*

Der Standardfehler einer Differenz beträgt

$$SED = \sqrt{\frac{2\sigma^2}{n}}$$

Diese Gleichung kann umgeformt werden nach *n*.

$$n = \frac{2\sigma^2}{SED^2}$$

Beispiel Ross and Knodt (1948)

- Fütterungsexperiment mit Schwarzbunten
- Tägliche Zunahmen in Pfund (lb)
- Vorversuch mit s^2 = 2199 lb.^{2}
- Mittelwerte in der Größenordnung 200 lb
- Geplanter Versuch: *SED* = 20 lb angestrebt

$$n = \frac{2\sigma^2}{SED^2} = \frac{2 \times 2199}{20^2} = 11$$

Zulässige Abweichung von wahrer Differenz

Strebe an:

$$P\left(\left|\bar{d}\right| < \tau_d\right) = 1 - \alpha$$

Wähle:

$$n = \frac{2\sigma^2 z_{1-\alpha/2}^2}{\tau_\delta^2}$$

τ_δ = Zulässige Abweichung der geschätzten Differenz $\bar{d}$ vom wahren Wert δ

$z_{1-\alpha/2}$ = $(1-\alpha/2)$-Quantil der Standardnormalverteilung

Beispiel Ross and Knodt (1948)

- Zulässige Abweichung τ_d = 20 lb
- α = 5% $\Rightarrow z_{1-\alpha/2} \approx 2$

$$n \approx \frac{8\sigma^2}{\tau_\delta^2} = \frac{8 \times 2199}{20^2} = 43.98 \Rightarrow 44$$

$\Rightarrow$ Vier mal so viel wie nötig für *SED* = 20 lb

$\Rightarrow$ Erreichte Genauigkeit hier *SED* = 10 lb

Das ganze rumgedreht

$$n = \frac{2\sigma^2 z_{1-\alpha/2}^2}{\tau_\delta^2}$$

$\Leftrightarrow$

$$\tau_\delta = \sqrt{\frac{2\sigma^2}{n}} z_{1-\alpha/2} = SED \times z_{1-\alpha/2} \approx 2 \times SED$$

$\Rightarrow$ Die mit gegebenem Design einhaltbare zulässige Abweichung beträgt 2 $\times$ *SED*

Erwartete halbe Breite eines 95%-Vertrauensintervalls

Vertrauensintervall für δ mit $(1-\alpha)\times 100\%$ Überdeckungswahrscheinlichkeit

$$\bar{d} \pm t_{w;1-\alpha/2}\sqrt{\frac{2s^2}{n}}$$

Halbe Breite

$$HW = t_{w;1-\alpha/2}\sqrt{\frac{2s^2}{n}}$$

---

w = Fehlerfreiheitsgrade

Approximation:

Erwartete halbe Breite des Vertrauensintervalls

$$EHW \approx z_{1-\alpha/2}\sqrt{\frac{2\sigma^2}{n}}$$

Auflösen nach *n*

$$n \approx \frac{2\sigma^2 z_{1-\alpha/2}^2}{EHW^2}$$

Beispiel Ross and Knodt (1948)

- Strebe an *EHW* = 20 lb
- α = 5% $\Rightarrow z_{1-\alpha/2} \approx 2$

$$n \approx \frac{8\sigma^2}{EHW^2} = \frac{8 \times 2199}{20^2} = 43.98 \Rightarrow 44$$

$\Rightarrow$ Vier mal so viel wie nötig für *SED* = 20 lb

$\Rightarrow$ Erreichte Genauigkeit hier *SED* = 10 lb

Das ganze rumgedreht

$$n = \frac{2\sigma^2 z_{1-\alpha/2}^2}{EHW^2}$$

$\Leftrightarrow$

$$EHW = \sqrt{\frac{2\sigma^2}{n}} z_{1-\alpha/2} = SED \times z_{1-\alpha/2} \approx 2 \times SED$$

$\Rightarrow$ Die mit gegebenem Design zu erwartende halbe Breite des VI beträgt 2 $\times$ *SED*

Vorgabe der LSD

$$LSD = t_{w;1-\alpha/2}\sqrt{\frac{2s^2}{n}} \approx z_{1-\alpha/2}\sqrt{\frac{2\sigma^2}{n}} = ELSD$$

Auflösen nach *n*

$$n \approx \frac{2\sigma^2 z_{1-\alpha/2}^2}{ELSD^2}$$

Beispiel Ross and Knodt (1948)

- Strebe an *LSD* = 20 lb
- α = 5% $\Rightarrow z_{1-\alpha/2} \approx 2$

$$n \approx \frac{8\sigma^2}{ELSD^2} = \frac{8 \times 2199}{20^2} = 43.98 \Rightarrow 44$$

$\Rightarrow$ Vier mal so viel wie nötig für *SED* = 20 lb

$\Rightarrow$ Erreichte Genauigkeit hier *SED* = 10 lb

Das ganze rumgedreht

$$n = \frac{2\sigma^2 z_{1-\alpha/2}^2}{ELSD}$$

$\Leftrightarrow$

$$ELSD = \sqrt{\frac{2\sigma^2}{n}} z_{1-\alpha/2} = SED \times z_{1-\alpha/2} \approx 2 \times SED$$

$\Rightarrow$ Die mit gegebenem Design zu erwartende LSD beträgt 2 × *SED*

Stichprobenplanung für Signifikanztest

$$n \approx \frac{2\sigma^2 \left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{\delta^2}$$

δ = kleinste relevante nachzuweisende Differenz

Tabelle: Werte von $C_{\alpha,\beta} = \left(z_{1-\alpha/2} + z_{1-\beta}\right)^2$ für typische Werte von α und β

		β			
		5%	10%	15%	20%
α	1%	17.8	14.9	13.0	11.7
	5%	13.0	10.5	9.0	7.8
	10%	10.8	8.6	7.3	6.2

Für α = 5% und β = 10% finden wir:

$$n \approx \frac{21\sigma^2}{\delta^2}$$

Beispiel Ross and Knodt (1948)

- Strebe Nachweis von $\delta = 20$ lb. an
- Für α = 5% und β = 10%

$$n \approx \frac{21\sigma^2}{\delta^2} = \frac{21 \times 2199}{20^2} = 115.4 \Rightarrow 116$$

Erreichte Präzision: *SED* $\approx$ 6 lb.

Das ganze rumgedreht

$$n \approx \frac{2\sigma^2 \left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{\delta^2}$$

$\Leftrightarrow$

$$\delta = \sqrt{\frac{2\sigma^2}{n}} \left(z_{1-\alpha/2} + z_{1-\beta}\right) = SED \times \left(z_{1-\alpha/2} + z_{1-\beta}\right)$$

Für α = 5% und β = 15%:

$$z_{1-\alpha/2} + z_{1-\beta} \approx 3 \quad \Rightarrow \quad \delta \approx 3 \times SED$$

$\Rightarrow$ Die mit gegebenem Design nachweisbare kleinste Differenz beträgt 3 $\times$ *SED*

Die 1-2-3 Regel

- Spezifikation der angestrebten Genauigkeit kann direkt mit 1 × *SED* erfolgen

- Die Grenzdifferenz (LSD) ist etwa 2 × *SED* /
  Die erwartete halbe Breite eines Vertrauensintervalls ist etwa 2 × *SED* /
  Etwa 95% der geschätzten Differenzen weicht um höchstens 2 × *SED* von der wahren Differenz ab

- Die kleinste relevante Differenz, die mit einer Teststärke von 85% nachgewiesen werden kann, beträgt 3 × *SED*

Verallgemeinerung für lineare gemischte Modelle mit Statistikpaket

- Vorgabe von *SED*
- Dummy-Datensatz für geplantes Design (Stichprobenumfang)
- Analyse mit fixierter Varianz zur Berechnung von *SED*

$\Rightarrow$ Vergrößere Design solange bis *SED* erreicht

(Stroup 2002)

Beispiel Ross and Knodt (1948)

- Fütterungsexperiment mit Schwarzbunten
- Tägliche Zunahmen in Pfund (lb)
- Vorversuch mit s^2 = 2199 lb.^{2}
- Mittelwerte in der Größenordnung 200 lb
- Geplanter Versuch: *SED* = 20 lb angesprebt

```
data sim;
n=11;
do trt=1 to 2;
  do i=1 to n;
    y=1;
   output;
  end;
end;
run;

proc mixed data=sim;
class trt;
model y=trt;
lsmeans trt/diff;
parms (2199)/hold=1;
run;
```

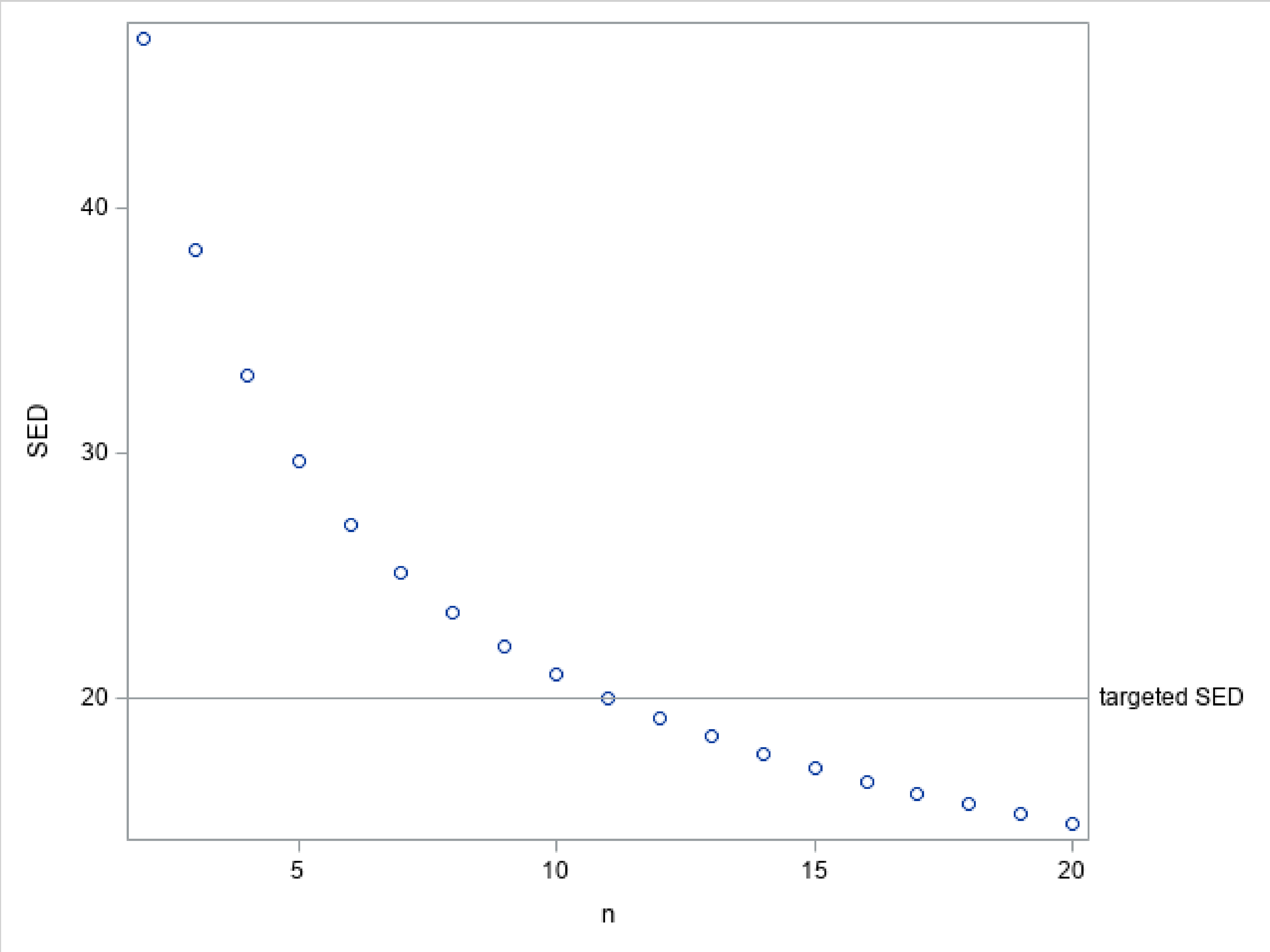
Differences of Least Squares Means

Effect	trt	_trt	Estimate	Standard Error	DF	t Value	Pr > |t|
trt	1	2	3.47E-16	19.9955	20	0.00	1.0000

Obs	n	trt	i	y
1	11	1	1	1
2	11	1	2	1
3	11	1	3	1
4	11	1	4	1
5	11	1	5	1
6	11	1	6	1
7	11	1	7	1
8	11	1	8	1
9	11	1	9	1
10	11	1	10	1
11	11	1	11	1
12	11	2	1	1
13	11	2	2	1
14	11	2	3	1
15	11	2	4	1
16	11	2	5	1
17	11	2	6	1
18	11	2	7	1
19	11	2	8	1
20	11	2	9	1
21	11	2	10	1
22	11	2	11	1

```
data sim;
do n=2 to 20;
  do trt=1 to 2;
    do i=1 to n;
      y=1;
     output;
    end;
  end;
end;
run;

proc mixed data=sim;
ods output diffs=diffs;
class trt;
model y=trt;
lsmeans trt/diff;
parms (2199)/hold=1;
by n;
run;
```


SED
40
30
20
targeted SED
5
10
15
20
n

Simulationsansatz für generalisierte lineare gemischte Modelle

- Simuliere *s* Datensätze für geplantes Design
- Verwende dafür das für die Auswertung vorgesehene Modell
- Brauche Vorinformation für Varianzen und feste Effekte
- Teststärke: Setze Treatmenteffekte so, dass angestrebtes δ für kleinste relevante Differenz realisiert wird
- Werte die *s* simulierten Datensätze aus im Hinblick auf *SED*, Teststärke, etc.

(Gbur et al. 2012; Green and McLeod 2015)