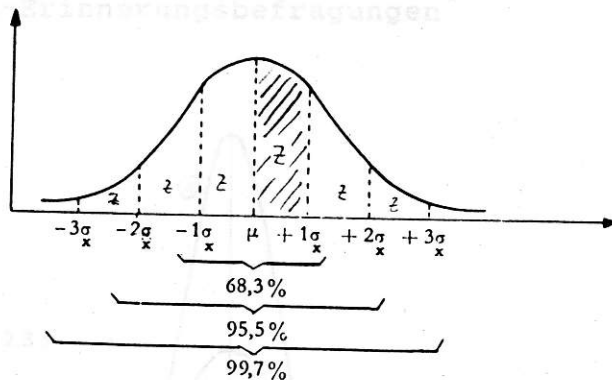


Abb. 60 Beispiele für die Berechnung des Stichprobenfehlers
 a) bei Normalverteilung der Meßwerte

- allgemeine Normalverteilungskurve



dabei ist μ der Mittelwert (mean) und

σ_x die Standardabweichung (SD = standard deviation).

Der Wert des Stichprobenfehlers $\hat{\sigma}_x$ ergibt sich aus der Formel:

$$\hat{\sigma}_x = \sqrt{\frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}}$$

wobei σ = Streuung der Grundgesamtheit

N = Umfang der Grundgesamtheit

n = Umfang der Stichprobe (Erhebungseinheiten)

Die zweite Wurzel $\sqrt{(N-n)/(N-1)}$ kann dann vernachlässigt werden, wenn der Anteil der Stichprobe an einer großen Grundgesamtheit klein ist (unter 5%), dann geht der Wert gegen 1.

Man erkennt, daß man im Prinzip die Streuung σ (SD) der Grundgesamtheit kennen muß, um den Stichprobenfehler $\hat{\sigma}_x$ zu berechnen, hat man diese nicht - wie in den meisten Fällen von ernährungsepidemiologischen Studien - dann muß man sie schätzen. In den Fällen von einer größeren Stichprobe ($n > 30$) nimmt man die Standardabweichung der Stichprobe dafür. So wird

$$\hat{\sigma}_x = \sqrt{\frac{s^2}{n}} \quad \text{wobei } s^2 = \frac{\sum (x - \bar{x})^2}{n}$$

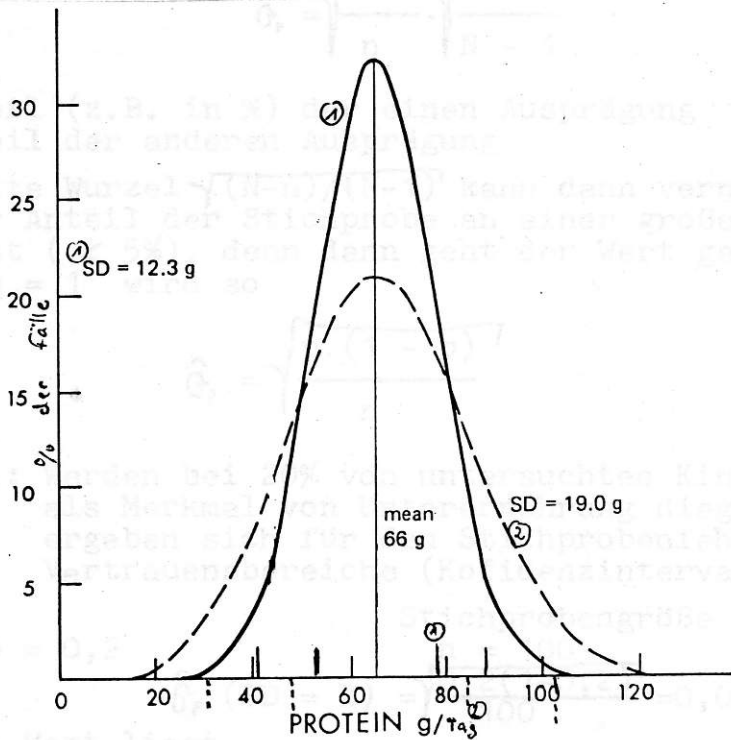
So wird also aus der Standardabweichung (SD) bei einfachen Zufallsstichproben der Standardfehler (SE = standard error), beide Werte sind identisch, wobei der Standardfehler in seinem Sinngehalt den Vertrauensbereich (Vertrauensintervall) angibt. Der wahre Wert liegt mit bestimmter Wahrscheinlichkeit um das gefundene Mittel (mean), wobei Stichprobenfehlerbreiten als Z-Werte angewandt werden ($Z = 1$ Standardfehler (SE)).
 Dazu ein Beispiel aus der Praxis:

Subcommittee on Criteria for Dietary Evaluation, National Research Council, Nutrient Adequacy - Assessment using Food...

Dazu ein Beispiel aus der Praxis:

Verteilung der täglichen Eiweißaufnahme bei amerikanischen Erwachsenen, ermittelt nach

- 1 7-Tage-Protokollen
- 2 24-Stunden-Erinnerungsbefragungen *



Danach liegt der wahre Wert für die tägliche Eiweißaufnahme in folgenden Werte-Bereichen um das gefundene Mittel (mean) von 66g

	im Fall 1	im Fall 2
mit 68,3%iger Wahrscheinlichkeit ($\pm 1SD = \pm 1Z$)	53,7 - 78,3	47 - 85 g Eiweiß/Tag
mit 95,5%iger Wahrscheinlichkeit ($\pm 2SD = \pm 2Z$)	41,4 - 90,6	28 - 104 g Eiweiß/Tag
mit 99,7%iger Wahrscheinlichkeit ($\pm 3SD = \pm 3Z$)	29,1 - 102,9	9 - 123 g Eiweiß/Tag

Man erkennt deutlich, daß die Vertrauensbereiche (die "wahren" Werte) je nach Verteilung und gewünschter Sicherheit der Aussage (Wahrscheinlichkeit) sehr unterschiedlich sind. (Vertrauensbereich = Kofidenz-Intervall, confidence intervals)

* Subcommittee on Criteria for Dietary Evaluation. National Research Council: Nutrient Adequacy - Assessment using Food Consumption Surveys. National Academy Press, Washington, DC, 1986 (selbst)

Abb. 60b) bei dichotomer Ausprägung des Merkmales

Die Formel für den Stichprobenfehler $\hat{\sigma}_p$ bei dichotomen Merkmalsausprägungen, d.h. das Merkmal kommt nur in zwei verschiedenen Möglichkeiten vor (z.B. Mangelsymptome für Fehlernährung; vorhanden/nicht vorhanden):

$$\hat{\sigma}_p = \sqrt{\frac{p \cdot q}{n}} \cdot \sqrt{\frac{N - n}{N - 1}}$$

p = Anteil (z.B. in %) der einen Ausprägung N = Grundgesamtheit
 q = Anteil der anderen Ausprägung n = Stichproben-Größe

Die zweite Wurzel $\sqrt{(N-n)/(N-1)}$ kann dann vernachlässigt werden, wenn der Anteil der Stichprobe an einer großen Grundgesamtheit klein ist (< 5%), denn dann geht der Wert gegen 1.

Da p + q = 1 wird so

$$\hat{\sigma}_p = \sqrt{\frac{p \cdot (1 - p)}{n}}$$

Beispiel: Werden bei 20% von untersuchten Kindern Muskelschwund als Merkmal von Unterernährung diagnostiziert, so ergeben sich für den Stichprobenfehler bzw. die Vertrauensbereiche (Kofidenzintervalle) folgende Werte

	Stichprobengröße	
	n = 100	n = 1600
p = 0,2	$\hat{\sigma}_p$ (SD = Z) = $\sqrt{\frac{0,2(1-0,2)}{100}} = 0,04$	$\sqrt{\frac{(0,2)(1-0,2)}{1600}} = 0,01$
der wahre Wert liegt mit 68,3%iger Wahrscheinlichkeit ($\pm 1\hat{\sigma}_p$)	0,16 - 0,24 16 - 24%	0,19 - 0,21 19 - 21%
mit 95,5%iger Wahrscheinlichkeit ($\pm 2\hat{\sigma}_p$)	0,12 - 0,28 12 - 28%	0,18 - 0,22 18 - 22%
mit 99,7%iger Wahrscheinlichkeit ($\pm 3\hat{\sigma}_p$)	0,08 - 0,32 8 - 32%	0,17 - 0,23 17 - 23%