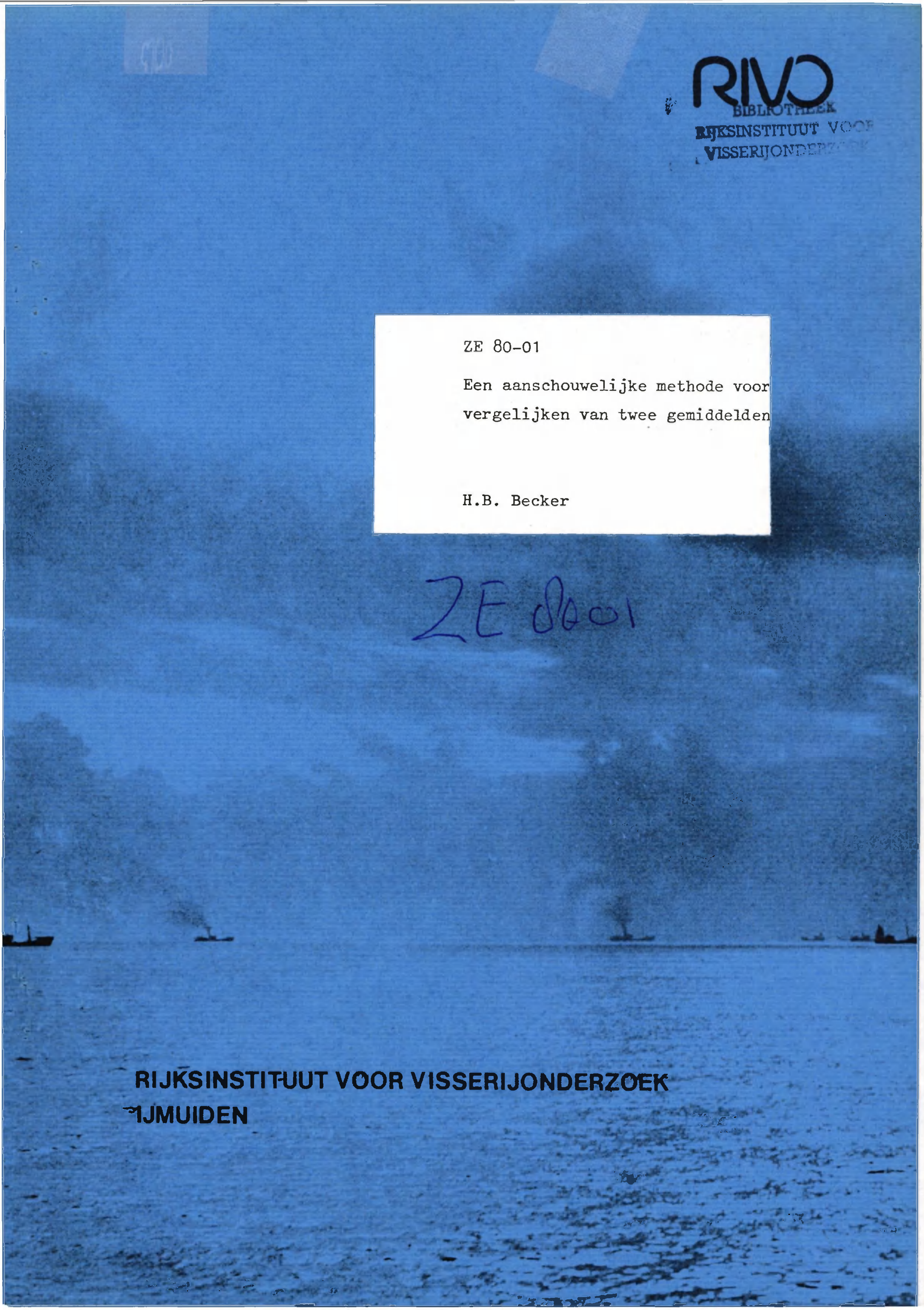


ZE 80-01

Een aanschouwelijke methode voor vergelijken van twee gemiddelden

H.B. Becker

RIJKSINSTITUUT VOOR VISSERIJONDERZOEK
IJMUIDEN

RIJKSINSTITUUT VOOR VISSERIJONDERZOEK

Haringkade 1 – Postbus 68 – IJmuiden – Tel. (02550) 1 91 31

Afdeling: Biologisch Onderzoek Zoutwatervisserij

Rapport: ZE 80-01

Een aanschouwelijke methode voor het vergelijken van twee gemiddelden.

Auteur: H.B. Becker

Project:

Projectleider:

Datum van verschijnen:

Inhoud:

I. Inleiding

II. Methode

III. Resultaten

EEN AANSCHOUWELIJKE METHODE VOOR HET VERGELIJKEN VAN TWEE GEMIDDELDEN

I. INLEIDING.

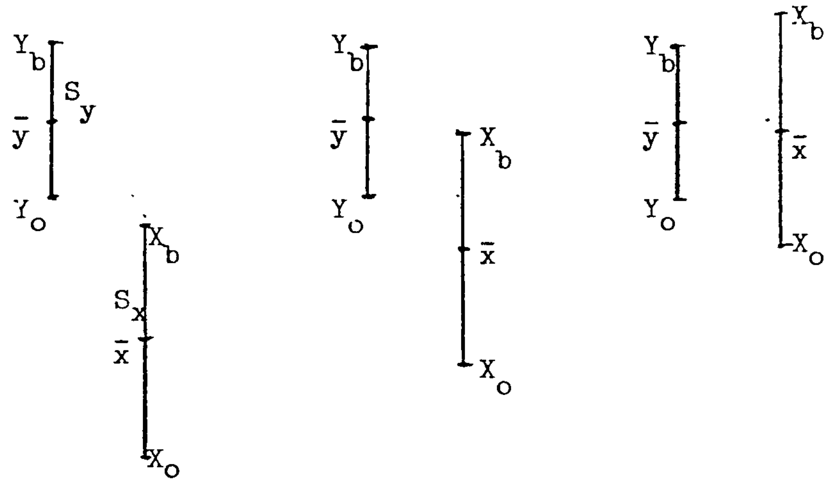
Veronderstel dat men beschikt over een aselecte steekproef van n waarnemingen $x_1,\ldots, x_n$ uit een normale $N(\mu_x,\delta_x^2)$ - verdeling en een aselecte steekproef van m waarnemingen $y_1,\ldots,y_m$ uit een normale $N(\mu_y,\delta_y^2)$ - verdeling. We willen nu op grond van de informatie uit de twee steekproeven gaan toetsen of μ_x en μ_y significant verschillend zijn. In het visserijonderzoek is dit een veel voorkomend probleem. Indien aangenomen kan worden dat de varianties δ_x^2 en δ_y^2 gelijk zijn, kunnen we voor het toetsen van de hypothese $\mu_x=\mu_y$ Student's t-toets voor twee steekproeven gebruiken. Wordt met behulp van de F-toets de hypothese van gelijkheid van varianties verworpen, dan kunnen we alsnog via een benaderingsmethode de significatie van het verschil van de gemiddelden toetsen.

Een andere methode, waarmee we ons in dit verhaal verder zullen bezighouden, is op een meer directe aanschouwelijke wijze trachten iets te zeggen over het significant verschillend zijn van μ_x en μ_y. Daartoe zullen we de informatie uit een steekproef in een figuur samenvatten als steekproefgemiddelde $\pm$ standaard-afwijking. Aangetoond zal worden dat de mate van overlapping of scheiding tussen de intervallen ons iets zal kunnen vertellen over het al dan niet significant verschillend zijn van de twee gemiddelden, mits rekening wordt gehouden met de relatieve lengten van de intervallen en de steekproefomvang.

II. METHODE.

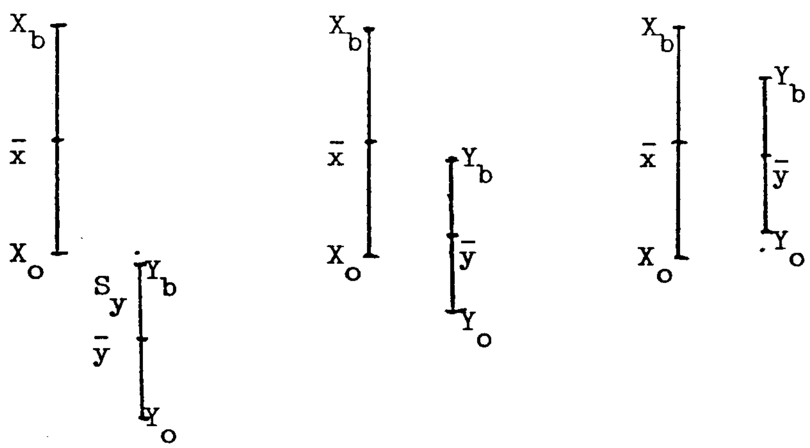
Het steekproefgemiddelde van beide steekproeven noteren we als $\bar{x}$ en $\bar{y}$ en de bijbehorende standaardafwijking als s_x en s_y. De intervalschattingen van μ_x en μ_y zijn respectievelijk $\bar{x} \pm s_x$ en $\bar{y} \pm s_y$. Zij M het quotient van de lengten van beide intervallen, dus $M=s_x/s_y$. We nemen verder aan dat $M \geq 1$. In fig. 1 en 2 zijn de verschillende mogelijkheden voor de twee intervallen weergegeven, waarbij onderscheid is gemaakt tussen $\bar{x} > \bar{y}$ en $\bar{x} \leq \bar{y}$.

fig. 1. Drie verschillende mogelijkheden bij het vergelijken van intervalschattingen in geval $\bar{x} \leq \bar{y}$



(a) (b) (c)

fig. 2. Drie verschillende mogelijkheden bij het vergelijken van intervalschattingen in geval $\bar{x} > \bar{y}$



(a) (b) (c)

We introduceren vervolgens de grootheid D gedefinieerd als volgt:

$$D = \frac{Y_b - X_b}{Y_b - Y_o} = \frac{(\bar{y} - \bar{x}) - (s_x - s_y)}{2s_y} \qquad \text{als } \bar{x} \leq \bar{y}$$

$$D = \frac{X_o - Y_o}{Y_b - Y_o} = \frac{(\bar{x} - \bar{y}) - (s_x - s_y)}{2s_y} \qquad \text{als } \bar{x} > \bar{y}$$

M.a.w.,

$$D = \frac{|\bar{x} - \bar{y}| - (s_x - s_y)}{2s_y}$$

waarbij $|\bar{x} - \bar{y}| = \bar{x} - \bar{y}$ als $\bar{x} > \bar{y}$

$\qquad\qquad\qquad = \bar{x} - \bar{y}$ als $\bar{x} \leq \bar{y}$

Met behulp van de definitie van D en fig.1 en 2 is nu gemakkelijk na te gaan dat geldt:

— intervallen overlappen elkaar niet (fig. 1_a en 2_a) : $D > 1$
  intervallen raken elkaar: $D = 1$

— intervallen overlappen elkaar gedeeltelijk (1_b en 2_b) : $D < 1$

— grootste interval overlapt kleinste interval geheel
  (fig. 1_c en 2_c) : $D \leq 0$

De minimale waarde van D wordt bereikt als $\bar{x} = \bar{y}$; in dat geval geldt:

$$D = \frac{-(s_x - s_y)}{2s_y} = \frac{-(M - 1)}{2}$$

We kunnen ons nu gaan afvragen bij welke mate van overlapping de gemiddelden μ_x en μ_y nog significant verschillend zijn. Bij aanname van gelijkheid van varianties kijken we daartoe naar de gebruikelijke toetsgrootheid

$$T = \frac{|\bar{x} - \bar{y}|}{s(^1/n + {}^1/m)^{\frac{1}{2}}}$$

waarbij

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2 + \Sigma(y_i - \bar{y})^2}{n + m - 2}$$

Deze toetsgrootheid bezit de Student's t_{n+m-2}-verdeling. De gelijkheid van varianties dient echter eerst te worden getoetst via de gebruikelijke toetsgrootheid

$$F = \frac{s_x^2}{s_y^2} = M^2$$

welke de $F_{n-1,m-1}$-verdeling bezit.

Indien echter met behulp van de F-toets de hypothese van gelijkheid van varianties moet worden verworpen, kan men gebruik maken van een benaderde grootheid voor het toetsen van een significant verschil tussen de gemiddelden μ_x en μ_y, nl.

$$T' = \frac{|\bar{x} - \bar{y}|}{\left(\frac{s_x^2}{n} + \frac{s_y^2}{m} \right)^{\frac{1}{2}}}$$

welke de Student's t-verdeling bezit met

$$\frac{(1 + V)^2}{\frac{V^2}{n+1} + \frac{1}{m+1}} - 2 \qquad \text{vrijheidsgraden}$$

waarbij $V = \frac{M^2 m}{n}$

Substitueren we nu de uitdrukking voor $|\bar{x} - \bar{y}|$ als functie van de toetsgrootheid in de formule voor D, dan blijkt dat D een functie is van T (of T'), M, m en n. Verder zien we dat D niet afhankelijk is van de absolute lengte van de intervallen afzonderlijk doch slechts van hun quotient.

Nemen we nu voor T en T' de kritieke waarden van de t-verdeling op 5% niveau, m.a.w. de kleinste waarden voor T en T', zodanig dat μ_x en μ_y nog juist significant verschillend zijn op 5% niveau, dan vinden we op deze wijze de kleinste waarden van D, zodanig dat nog juist significantie optreedt op 5% niveau.

Deze "kleinste" waarden van D zijn voor verschillende waarden van M, n en m (T en T' zijn door m en n bepaald) betrekkelijk gemakkelijk te berekenen.

In tabel 1 zijn enkele berekende minimumwaarden van D weergegeven.

Tabel 1. Minimumwaarden van D, waarbij nog juist significantie optreedt.

n = 10 m \ M	1	1.5	2	2.5	3
5	0.59	0.56	0.54	0.52	0.13
10	0.47	0.36	0.24	0.19	0.10
15	0.42	0.27	0.24	0.17	0.10
20	0.40	0.23	0.24	0.17	0.09
25	0.38	0.19	0.23	0.16	0.10

III. RESULTATEN.

Er blijken betrekkelijk weinig algemene regels te bestaan bij aanschouwelijk onderzoek naar significantie tussen twee gemiddelden. Voor $n \geq 10$ zijn de minimumwaarden van D waarbij de gemiddelden nog juist significant verschillend zijn, alle kleiner dan 1. Dat betekent dus dat voor $n \geq 10$ en geen overlapping van de intervallen ($D > 1$), er significantie optreedt.
Ook treedt er significantie op als $n \geq 10$ en meer dan 2/3 van het kleinste interval niet bedekt is ($D > 0.67$). Verder zijn voor $M \geq 3.0$ en $n \geq 15$ de gemiddelden significant verschillend wanneer het kleinste interval niet geheel bedekt wordt door het grootste interval. ($D > 0$).
In het algemeen blijken dit soort vuistregels bij steekproefgroottes kleiner dan zes erg onbetrouwbaar te zijn.