KM 30(1989)
pag 47 - 55



# EEN TOEPASSING VAN EEN MODEL VOOR OVER- EN/OF ONDERDISPERSIE BY BINOMIALE VARIATIE

Bas Engel

## 0 SAMENVATTING

Dit artikel geeft een toepassing van een model voor over- en/of onderdispersie ten opzichte van de variantie van een Binomiale verdeling. Parameterschattingen volgen uit een combinatie van maximum quasi-likelihood (McCullagh en Nelder, 1983) voor de locatie-parameters en een momentenmethode voor de over-/onderdispersie parameter (Williams, 1982). Voor de numerieke verwerking is het statistische pakket Genstat (Genstat 5, 1987) gebruikt.

---

Groep Landbouwwiskunde
Postbus 100
6700 AC Wageningen
tel.: 08370-19100

Met dank aan Ir. Joop te Brake (voor het beschikbaar stellen van de gegevens) en Willem Buist (voor assistentie bij de numerieke verwerking), beide medewerkers van het IVO 'Schoonoord' te Zeist.

## 1 DE PROEFOPZET

Voor een onderzoek naar de effecten van aktieve immunisatie tegen Androsteendion op de vruchtbaarheid van Texelaar ooien, uitgevoerd op het Instituut voor Veeteeltkundig Onderzoek "Schoonoord" te Zeist, zijn 125 ooien verdeeld in een controlegroep en een proefgroep. De dieren in de proefgroep zijn behandeld met het preparaat Fecundin. Van deze behandeling wordt na dekking een stijging van het aantal ovulaties en van het aantal embryos verwacht.

De dieren zijn ingedeeld in 4 leeftijdsklassen: $\leq {}^1/_2$ jaar, ${}^1/_2$-$1{}^1/_2$ jaar, $1{}^1/_2$-$2{}^1/_2$ jaar en $> 2{}^1/_2$ jaar. Verder is er sprake van 2 dekperioden. Er zijn dus 3 proeffactoren: behandeling, leeftijd en dekperiode met respectievelijk 2, 4 en 2 niveaus.

In dit artikel kijken we naar de analyse van het aantal embryos, conditioneel op het aantal ovulaties. De waarnemingen zijn gegeven in tabel A1 in de appendix. Vier dieren, die niet drachtig werden (waarvoor het aantal ovulaties dus gelijk aan nul is) en twee dieren, waarvan het aantal ovulaties niet bekend was, zijn niet in de analyse opgenomen.

## 2 HET MODEL

Stel voor een willekeurige ooi:
- $n$ = het aantal ovulaties,
- $x$ = het aantal embryos,
- $p$ = de kans dat een willekeurige ovulatie in een embryo resulteert.

Voor ieder van de ovulaties voeren we een indicator variabele in:
$a_m$ = 1 als de m-de ovulatie in een embryo resulteert
en 0 als dit niet het geval is, $m=1,\ldots,n$.

Nu geldt dus $x = \Sigma_{m=1}^{n} a_m$.

De kans p hangt af van de behandeling, leeftijdsklasse en dekperiode, èn van het dier zelf.

$$p = p_0 + \varepsilon.$$

Hierin is $p_0$ bepaald door de proeffactoren en is $\varepsilon$ de stochastische bijdrage van het dier. We nemen aan dat $\varepsilon$ gemiddelde 0 heeft en een variantie van de vorm:

$$\mathrm{Var}(\varepsilon) = \sigma^2 p_0(1-p_0).$$

Voor $p_0$ in de buurt van 0 of 1 is de tussen-dier variatie dus klein en voor $p_0$ rond 0.5 is de tussen-dier variatie groot.

We veronderstellen dat er sprake kan zijn van een zekere 'competitie' tussen de ovulaties. We voeren daartoe een correlatie $\rho$ in tussen $a_m$ en $a_{m'}$, conditioneel op de dierbijdrage $\varepsilon$:

$$\mathrm{cov}(a_m, a_{m'} \mid \varepsilon) = \rho\, p(1-p) \qquad m, m' = 1, \ldots, n, \quad m \neq m'.$$

We denken daarbij aan negatieve waarden voor $\rho$. Aangenomen wordt dat $\rho$ in absolute waarde niet al te groot is. Op z'n minst zal moeten gelden:

$$-1/(n_{max}-1) < \rho < 1,$$

waarin $n_{max}$ het grootste aantal ovulaties is dat kan voorkomen. De correlatie $\rho$ wordt verder constant verondersteld.

De kans $p_0$, het 'fixed' gedeelte van p, kan eventueel van n afhangen. Om het model niet te complex te maken is hiervoor een sterk vereenvoudigde voorziening getroffen. Naast de eerder geïntroduceerde factoren behandeling (B, niveau 1 is de proefgroep en niveau 2 de controlegroep), leeftijd (L) en dekperiode (D) voeren we een factor A (van 'aantal') in met 2 niveaus:

$A = 1$ wanneer $n \leq 2$ en 2 wanneer $n > 2$.

De kans $p_0$ hangt nu van B, L, D èn A af.

Voor het modelleren van $p_0$ gebruiken we de logit-transformatie:

$$\mathrm{logit}\,(p_0) = \log\left(\frac{p_0}{1-p_0}\right) = \text{algemeen gemiddelde + som van hoofdeffecten en interakties voor de factoren B, L, D en A.}$$

Om een onevenredig groot aantal parameters te vermijden zijn geen interakties met factor A opgenomen.
Voor de verwachting en variantie van het aantal embryos x geldt nu:

$$E(x) = np_0,$$
$$\mathrm{Var}(x) = np_0(1-p_0)\{1 + (n-1)\phi\},$$

waarin $\phi = \rho + (1-\rho)\sigma^2$, de (niet-conditionele) correlatie tussen $a_m$ en $a_{m'}$.

De variantie is afgeleid in appendix A2. Merk op dat de correlatie $\phi$ zowel positief als negatief kan zijn. Deze dispersie parameter is het totaal van overdispersie geïntroduceerd door de dierbijdrage $\varepsilon$ (tussen-dier variatie) en onderdispersie geïntroduceerd door de (conditionele) 'competitie' correlatie $\rho$ (binnen-dier variatie).

## 3 DE ANALYSE

Stel dat $\phi$ bekend is.
Het verband tussen E(x) en Var(x) ligt nu vast. Dit maakt het mogelijk om de onbekende parameters in logit($p_0$) te schatten met de Quasi Likelihood methode (QL). QL is beschreven in McCullagh en Nelder (1983). De methode is beknopt weergegeven in appendix A3. De praktische uitvoering verloopt probleemloos met het statistische pakket Genstat (Genstat 5, 1987).

Stel dat de vector van parameters $\underline{\beta}$ is gepartitioneerd als $\underline{\beta} = (\underline{\beta}_1', \underline{\beta}_2')'$. Voor een toets van de hypothese $H_0$: $\underline{\beta}_2 = \underline{0}$, kan de Quasi Likelihood Ratio (QLR) worden gebruikt. Onder $H_0$ volgt $-2 \log(QLR)$ bij benadering een chikwadraat verdeling met als vrijheidsgraden het aantal vrije parameters in $\underline{\beta}_2$ (de reductie in dimensie door het opleggen van de nulhypothese), zie hiervoor McCullagh (1983).

Bij het gebruik van Genstat volgt $-2 \log(QLR)$ uit het verschil van de zogenaamde deviances onder het gerestricteerde model ($\underline{\beta}_2 = \underline{0}$) en het ongerestricteerde model.
Een alternatief is de Quasi Wald-toets gebaseerd op de asymptotische normaliteit van de schatters op de logit-schaal.

Uiteraard is $\phi$ niet bekend. We gebruiken daarom in het voorafgaande een schatting $\tilde{\phi}$. Deze schatting kan worden afgeleid volgens een momenten-methode op basis van de grootheid van Pearson die is gesuggereerd door Williams (1982). De schattingsmethode is iteratief, beschreven in appendix A4 en eenvoudig in Genstat te implementeren. Het kan ook met GLIM, zie hiervoor Williams (1982).

## 4 DE RESULTATEN

Het grootste model dat is bekeken is:

$$\text{logit}\,(p_0) = \mu + B + L + D + A + B.L + B.D + L.D + B.L.D\ .$$

De schatting voor $\phi$ in dit model is: $\tilde{\phi} = -0.0751$. De negatieve uitkomst is een indicatie dat binnen-dier 'competitie' inderdaad een rol zou kunnen spelen. In onderstaande tabel zijn de interakties getoetst m.b.v. de QLR.

| Interaktie | $-2 \log (QLR)$ | vrijheidsgraden | overschrijdingskans |
|---|---|---|---|
| B.L.D | 3.9 | 3 | 0.27 |
| B.L. | 1.1 | 3 | 0.78 |
| B.D. | 2.1 | 1 | 0.15 |
| L.D | 1.3 | 3 | 0.73 |

Voor bijvoorbeeld de twee-factor interaktie L.D gebruiken we de deviances van de volgende twee modellen:

$$\mu + B + L + D + A + B.L + B.D + L.D$$

en

$$\mu + B + L + D + A + B.L + B.D$$

Naar aanleiding van de toetsingsresultaten reduceren we het model tot alleen de hoofdeffecten:

$$\text{logit}\,(p_0) = \mu + B + L + D + A.$$

In dit model bedragen de overschrijdingskansen voor B, L, D en A achtereenvolgens 0.06, 0.14, 1.0 en 0.07. We concluderen dat er geen aantoonbaar verschil bestaat tussen de dekperioden. Kijken we voor de factor leeftijd specifiek naar het contrast tussen de klassen 1 en 4 (de jongste en de oudste dieren) dan is de overschrijdingskans 0.04. Leeftijd is dus van belang.

Zonder factor A in het model draagt de factor behandeling zeer significant bij ($p < 0.01$). Globaal ligt het quotient x/n met $n \geq 1$ voor de behandelde dieren 0.12 lager dan voor de controle dieren. Een rangtoets op de quotienten x/n (waarin nu ook n als stochastisch wordt beschouwd) geeft een zeer significant resultaat ($p < 0.01$). Deze rangtoets is beknopt weergegeven in appendix A5. Dit is geen bewijs, maar wel een sterke aanwijzing voor de juistheid van de volgende bewering:

> Voor de behandelde dieren is de embryonale ontwikkeling significant lager dan voor de controle dieren. Echter, dit is voor een deel terug te voeren tot de veel lagere kans op ontwikkeling die gepaard gaat met hogere aantallen ovulaties.

Wanneer factor A niet in het model wordt opgenomen is er geen aantoonbaar leeftijdseffect. Mogelijk werken hier twee krachten tegen elkaar in:

(i) een hogere leeftijdsklasse geeft een hoger aantal ovulaties, maar bij een hoger aantal ovulaties is de ontwikkelingskans geringer;

(ii) een hogere leeftijd geeft bij een gegeven aantal ovulaties een hogere kans op ontwikkeling.

Zonder A in het model heffen (i) en (ii) elkaar grotendeels op. Nemen we A op, dan wordt voor (i) (voor een groot deel) gecorrigeerd en komt (ii) significant te voorschijn.

De parameterschattingen op logit-schaal zijn weergegeven in onderstaande tabel.

| Parameter | Schatting | Standaard-afwijking |
|---|---|---|
| $\mu$ | 1.41 | 0.18 |
| $B_1$ | -0.33 | 0.18 |
| $B_2$ | 0.33 | 0.18 |
| $L_1$ | -0.55 | 0.35 |
| $L_2$ | -0.20 | 0.23 |
| $L_3$ | 0.18 | 0.28 |
| $L_4$ | 0.58 | 0.29 |
| $D_1$ | -0.03 | 0.15 |
| $D_2$ | 0.03 | 0.15 |
| $A_1$ | 0.33 | 0.18 |
| $A_2$ | -0.33 | 0.18 |

Schatten we $\phi$ opnieuw in het hoofdeffectenmodel dan vinden we $\tilde{\phi} = -0.1175$. De schattingen veranderen weinig, de overschrijdingskansen worden iets kleiner en liggen voor A en B net boven de 0.05.

## 5. SLOTOPMERKINGEN

Er zijn aparte analyses uitgevoerd op het aantal ovulaties en het aantal embryos per dier, deze komen in dit verhaal verder niet aan de orde. In principe is het mogelijk één model te construeren voor ovulaties en embryos samen. De omvang van de dataset in combinatie met het discrete karakter van de waarnemingen maakt dit echter weinig aantrekkelijk.

Voor de combinatie van quasi-likelihood en schatting van de dispersie parameter $\phi$ aan de hand van Pearsons grootheid is in de literatuur nog geen theoretische verantwoording gegeven.

## 6. REFERENTIES

Conover, M.J. "Practical non-parametric statistics". 2nd edition, John Wiley & Sons, 1980.

Genstat 5 Committee, Payne R.W. en anderen. "Genstat 5 Reference Manual", Clarendon Press - Oxford, 1987.

McCullagh, P. "Quasi-likelihood functions". The Annals of Statistics 11, pp. 59-67, 1983.

McCullagh, P., Nelder, J.A. "Generalized linear models". Chapman and Hall, 1983.

Searle, S.R. "Linear Models". John Wiley & Sons, 1971.

Williams, D.A. "Extra-binomial variation in logistic linear models". Applied Statistics 31, pp. 144-148, 1982.

## APPENDICES

A1 Factoren : Behandeling, Leeftijd, Dekperiode
Variabelen : n(ovulaties), x(embryos)

| B | L | D | n | x |
|---|---|---|---|---|
| 1 | 1 | 1 | 2 | 1 |
| 1 | 1 | 1 | 2 | 2 |
| 1 | 1 | 1 | 2 | 2 |
| 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 2 | 1 |
| 1 | 1 | 1 | 2 | 2 |
| 1 | 1 | 2 | 2 | 1 |
| 1 | 1 | 2 | 2 | 1 |
| 1 | 1 | 2 | 1 | 1 |
| 1 | 1 | 2 | 2 | 1 |
| 1 | 1 | 2 | 2 | 1 |
| 1 | 1 | 2 | 1 | 1 |
| 1 | 2 | 1 | 2 | 1 |
| 1 | 2 | 1 | 2 | 2 |
| 1 | 2 | 1 | 2 | 2 |
| 1 | 2 | 1 | 4 | 2 |
| 1 | 2 | 1 | 3 | 2 |
| 1 | 2 | 1 | 2 | 2 |
| 1 | 2 | 1 | 2 | 2 |
| 1 | 2 | 1 | 3 | 1 |
| 1 | 2 | 2 | 2 | 2 |
| 1 | 2 | 2 | 2 | 1 |
| 1 | 2 | 2 | 2 | 2 |
| 1 | 2 | 2 | 2 | 1 |
| 1 | 2 | 2 | 3 | 2 |
| 1 | 2 | 2 | 2 | 2 |
| 1 | 2 | 2 | 3 | 3 |

| B | L | D | n | x |
|---|---|---|---|---|
| 1 | 2 | 2 | 3 | 2 |
| 1 | 2 | 2 | 3 | 1 |
| 1 | 2 | 2 | 3 | 2 |
| 1 | 2 | 2 | 2 | 1 |
| 1 | 3 | 1 | 3 | 3 |
| 1 | 3 | 1 | 3 | 2 |
| 1 | 3 | 1 | 2 | 1 |
| 1 | 3 | 1 | 3 | 3 |
| 1 | 3 | 1 | 2 | 2 |
| 1 | 3 | 1 | 2 | 1 |
| 1 | 3 | 2 | 4 | 4 |
| 1 | 3 | 2 | 4 | 3 |
| 1 | 3 | 2 | 4 | 1 |
| 1 | 3 | 2 | 3 | 3 |
| 1 | 3 | 2 | 3 | 2 |
| 1 | 3 | 2 | 2 | 2 |
| 1 | 4 | 1 | 3 | 2 |
| 1 | 4 | 1 | 2 | 1 |
| 1 | 4 | 1 | 3 | 2 |
| 1 | 4 | 1 | 3 | 3 |
| 1 | 4 | 1 | 2 | 2 |
| 1 | 4 | 1 | 3 | 3 |
| 1 | 4 | 2 | 2 | 2 |
| 1 | 4 | 2 | 4 | 4 |
| 1 | 4 | 2 | 2 | 2 |
| 1 | 4 | 2 | 4 | 3 |
| 1 | 4 | 2 | 2 | 2 |
| 1 | 4 | 2 | 2 | 2 |
| 1 | 4 | 2 | 5 | 2 |
| 2 | 1 | 1 | 1 | 1 |

| B | L | D | n | x |
|---|---|---|---|---|
| 2 | 1 | 1 | 2 | 1 |
| 2 | 1 | 1 | 2 | 1 |
| 2 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 1 | 1 |
| 2 | 1 | 2 | 1 | 1 |
| 2 | 1 | 2 | 1 | 1 |
| 2 | 1 | 2 | 1 | 1 |
| 2 | 1 | 2 | 1 | 1 |
| 2 | 1 | 2 | 1 | 1 |
| 2 | 1 | 2 | 1 | 1 |
| 2 | 2 | 1 | 2 | 2 |
| 2 | 2 | 1 | 1 | 1 |
| 2 | 2 | 1 | 3 | 2 |
| 2 | 2 | 1 | 2 | 1 |
| 2 | 2 | 1 | 1 | 1 |
| 2 | 2 | 1 | 2 | 1 |
| 2 | 2 | 1 | 2 | 2 |
| 2 | 2 | 1 | 2 | 1 |
| 2 | 2 | 1 | 2 | 2 |
| 2 | 2 | 1 | 3 | 2 |
| 2 | 2 | 1 | 2 | 2 |
| 2 | 2 | 2 | 2 | 1 |
| 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 1 | 1 |
| 2 | 2 | 2 | 2 | 2 |

| B | L | D | n | x |
|---|---|---|---|---|
| 2 | 2 | 2 | 1 | 1 |
| 2 | 2 | 2 | 1 | 1 |
| 2 | 3 | 1 | 2 | 2 |
| 2 | 3 | 1 | 3 | 3 |
| 2 | 3 | 1 | 2 | 1 |
| 2 | 3 | 1 | 2 | 2 |
| 2 | 3 | 1 | 3 | 2 |
| 2 | 3 | 1 | 2 | 2 |
| 2 | 3 | 2 | 2 | 1 |
| 2 | 3 | 2 | 2 | 2 |
| 2 | 3 | 2 | 2 | 1 |
| 2 | 3 | 2 | 2 | 2 |
| 2 | 3 | 2 | 2 | 2 |
| 2 | 3 | 2 | 2 | 2 |
| 2 | 4 | 1 | 2 | 1 |
| 2 | 4 | 1 | 2 | 2 |
| 2 | 4 | 1 | 2 | 2 |
| 2 | 4 | 1 | 2 | 2 |
| 2 | 4 | 1 | 2 | 2 |
| 2 | 4 | 1 | 3 | 3 |
| 2 | 4 | 1 | 2 | 2 |
| 2 | 4 | 1 | 2 | 2 |
| 2 | 4 | 2 | 2 | 2 |
| 2 | 4 | 2 | 2 | 2 |
| 2 | 4 | 2 | 2 | 1 |
| 2 | 4 | 2 | 2 | 2 |
| 2 | 4 | 2 | 2 | 2 |
| 2 | 4 | 2 | 3 | 3 |
| 2 | 4 | 2 | 3 | 3 |

A2 $\mathrm{Var}(x) = \mathrm{Var}\{E(x|\varepsilon)\} + E\{\mathrm{Var}(x|\varepsilon)\} =$
$= \mathrm{Var}(np) + E\{np(1-p) + n(n-1)\rho p(1-p)\} =$
$= n^2\mathrm{Var}(\varepsilon) + n\{1 + \rho(n-1)\}E\{p(1-p)\} =$
$= n^2\mathrm{Var}(\varepsilon) + n\{1 + \rho(n-1)\}\{p_0(1-p_0) - \mathrm{Var}(\varepsilon)\} =$
$= np_0(1-p_0)\{1 + \phi(n-1)\}$

A3 Stel dat $p_*$ een beginschatting is voor $p_0$.
$E(x) = np_0 \approx np_* + n\left[\frac{\delta p}{\delta \underline{\beta}}\right]'_{\underline{\beta}_*}(\underline{\beta}_0 - \underline{\beta}_*)$
$= np_* + np_*(1-p_*)\,\underline{b}'(\underline{\beta}_0 - \underline{\beta}_*),$

waarin $\underline{\beta}_*$ correspondeert met $p_*$ en $\underline{\beta}_0$ de echte waarde van $\underline{\beta}$ is met:

$\mathrm{logit}\,(p_0) = \underline{b}'\underline{\beta}_0.$

We voeren nu een nieuwe afhankelijke variabele in:

$Y = (x-np_*)/v_* + \underline{b}'\underline{\beta}_*$ met $v_* = np_*(1-p_*)$.

Nu geldt

$E(Y) \approx \underline{b}'\underline{\beta}_0$ en $\mathrm{var}(Y) \approx (wv_*)^{-1},$

waarin $w = \{1 + \phi(n-1)\}^{-1}$.

Met behulp van een gewogen lineaire regressie voor de variabele Y kunnen we nu een verbeterde schatting voor $\underline{\beta}_0$ afleiden. Deze noemen we weer $\underline{\beta}_*$ en we itereren tot het proces convergeert.

A4 Stel dat $\phi_*$ een beginschatting is voor de echte waarde $\phi_0$ van parameter $\phi$. Bepaal als in A3 een schatting $\tilde{\underline{\beta}}$ met bijbehorende $\tilde{\underline{p}}$.

Vat de gewichten w samen in een diagonaalmatrix $W = \mathrm{diag}(w)$, en voer de designmatrix X in waarvan de rijen de vectoren $\underline{b}$ zijn behorende bij de verschillende dieren.

Stel $W_*$ correspondeert met $\phi_*$ en $W_0$ met $\phi_0$.

Definieer Pearsons grootheid:

$$X^2_p = \Sigma\, w_* \frac{(x-n\tilde{p})^2}{n\tilde{p}(1-\tilde{p})}.$$

Nu geldt $X^2_p = (\underline{Y}_* - X\tilde{\underline{\beta}})'W_*V_*(\underline{Y}_* - X\tilde{\underline{\beta}})$,
waarin $\underline{Y}_*$ de variabele $\underline{Y}$ is geëvalueerd voor $\tilde{\underline{\beta}}$ en $\phi_*$
en $V_* = \mathrm{diag}(v_*)$ met $v_* = n\tilde{p}(1-\tilde{p})$.

$\tilde{\underline{\beta}} = QW_*V_*\underline{Y}_*$, waarin $Q = X(X'W_*V_*X)^{-1}X'$.

Vergeten we nu het stochastische karakter van $V_*$ dan volgt (zie Searle (1971), pag. 55):

$$E(X_p)^2 = \text{spoor}[E(X_p)^2] = E[\text{spoor}(X_p^2)] =$$
$$\approx \text{spoor}[(I-QW_*V_*)'W_*V_*(I-QW_*V_*)(W_0V_*)^{-1}] +$$
$$\underline{\beta}_0'X'(I-QW_*V_*)'W_*V_*(I-QW_*V_*)X\underline{\beta}_0.$$

Nu geldt $W_*V_*(I-QW_*V_*)X = 0$, dus

$$E(X_p^2) \approx \text{spoor}[(I-QW_*V_*)'W_*V_*(I-QW_*V_*)(W_*V_*)^{-1}W_*W_0^{-1}]$$
$$= \text{spoor}[(I-V_*W_*Q)W_*W_0^{-1}] =$$
$$= \Sigma\, w_*(1-v_*w_*q)/w_0$$

met q de diagonaal van Q.

Los nu een nieuwe waarde voor $\phi$ op uit:

$$X_p^2 = \Sigma w_*(1-v_*w_*q)(1+(n-1)\phi).$$

Itereer tot convergentie is bereikt. In dat geval geldt:

$X_p^2 = \text{spoor}[I-V_*W_*Q] = \text{rang}[I-V_*W_*Q] =$
$= N - \text{rang}(X)$, daar $V_*W_*Q$ idempotent is, N is het aantal dieren.
Het proces kan worden gestart met $\phi_* = 0$.

A5 Bepaal per dier x/n met $n \geq 1$.
Geef binnen iedere combinatie van factoren L en D rangnummers aan de quotienten. Laat T voor een combinatie de som van de rangnummers van de behandelde dieren zijn (dit is de toetsingsgrootheid van de toets van Wilcoxon). Onder de hypothese $H_0$: Fecundin heeft geen effect, geldt:

$$E(T) = a(b+1)/2, \quad \text{Var}(T) = \frac{a(b-a)}{b(b-1)}\,\Sigma R^2 - \frac{a(b-a)(b+1)^2}{4(b-1)},$$

waarin a het aantal behandelde dieren is, b het totaal aantal dieren en $\Sigma R^2$ de som van de kwadraten van de rangnummers van de behandelde dieren voor de betreffende combinatie, zie Conover (1980). Refereer $(\Sigma T - \Sigma\, E(T))/\sqrt{(\Sigma\, \text{Var}(T))}$ aan de standaard normale verdeling. Aangenomen is dat een eventueel effect van Fecundin consequent positief of negatief is voor de combinaties van L en D. Dit is redelijk te controleren met behulp van de toets van Wilcoxon uitgevoerd per combinatie van L en D. Alleen voor dekperiode 1 en leeftijdsklasse 1 bleek $T < E(T)$ (niet significant), voor de overige combinaties was $T > E(T)$.

Ontvangen: 02-09-1988
Geaccepteerd: 22-11-1988