Søren Johansen

A Note on the Welch-James Approximation to the Distribution of the Residual Sum of Squares in a Weighted Linear Regression

Preprint
March
1979
3

Institute of Mathematical Statistics
University of Copenhagen

Søren Johansen

A NOTE ON THE WELCH-JAMES APPROXIMATION TO THE DISTRIBUTION OF THE RESIDUAL SUM OF SQUARES IN A WEIGHTED LINEAR REGRESSION.

Preprint 1979 No. 3

INSTITUTE OF MATHEMATICAL STATISTICS

UNIVERSITY OF COPENHAGEN

March 1979

1. Summary

If in a regression problem the variances are not equal it is common to use the reciprocal estimated variances as weights. The residual sum of squares Q has asymptotically a χ^2 distribution when the degrees of freedom tend to infinity.

Welch (1947,1951) gave an approximation to the distribution of Q in the special case of the comparison of n means, by using a suitably chosen F distribution. James (1951,1954) gave an improved approximation using the fractiles of a χ^2 distribution and extended the results to the general linear model.

We shall show here how the results for the general linear model can be considerably simplified by using the technique due to Welch (1951), and extend the results to multivariate models and variance component models.

Results will also be given on the variance of the fitted value, thereby extending the results of Jacquez et al. (1968) to the general linear model.

2. Notation and main result

Let Y have an n-dimensional normal distribution with mean $\xi \in L_0$, a subspace of dimension $p < n$, and variance-covariance matrix $D = \text{diag}(\sigma_1^2,\ldots,\sigma_n^2)$. We shall assume that we have independent estimates $S = \text{diag}(s_1^2,\ldots,s_n^2)$ which have χ^2 distributions with degrees of freedom $(f_1,\ldots,f_n)$ and scale parameters $(\sigma_1^2/f_1,\ldots,\sigma_n^2/f_n)$.

We want to test the hypothesis $H : \xi \in L_1$, where $L_1 \subset L_0$ is a subspace of dimension $m < p$.

The weighted least squares estimate is found for $i = 0, 1$ by minimizing $(Y-\xi)'S^{-1}(Y-\xi)$ for $\xi \in L_i$, and is given by the projection $P_i(S)$ onto L_i with respect to S^{-1}.

As a test statistic for the hypothesis H we shall use the residual sum of squares

$$Q = Y'(P_0(S) - P_1(S))'S^{-1}(P_0(S) - P_1(S))Y$$
$$= Y'S^{-1}(P_0(S) - P_1(S))Y = Y'TY$$

say.

We can then formulate the main result.

<u>Theorem 1</u> Up to terms of order $1/f_i$ we have

(1) $$EQ = p - m + 2A + 2B$$

(2) $$VQ = 2(p - m) + 14A + 2B$$

where

(3) $$A = \Sigma_i (P_{0ii} - P_{1ii})(1 - P_{1ii})/f_i$$

(4) $$B = \Sigma_i (P_{0ii} - P_{1ii})(1 - P_{0ii})/f_i$$

Here $P_\nu = P_\nu(D)$, $\nu = 0,1$ and $P_{\nu ii}$ is the i'th diagonal element of the matrix P_ν.

We shall first comment on this result and then give some applications and defer the generalizations to section 4.

Notice first that for the case of testing a linear model in the full model we have $P_0 = I$ (the identity) and hence $B = 0$ and $A = \Sigma_i (1 - P_{1ii})^2/f_i$.

As a special case of this consider the problem of comparing n-means. Welch (1951) derived (1) and (2) with $B = 0$ and $A = \Sigma_i (1 - \sigma_i^{-2}/\Sigma_j \sigma_j^{-2})^2/f_i$. This result we can easily find since the projection P_1 in this case is given by

$$(P_1 Y)_i = \Sigma_j \sigma_j^{-2} Y_j / \Sigma_j \sigma_j^{-2}$$

from which we read off directly that $P_{1ii} = \sigma_i^{-2}/\Sigma_j \sigma_j^{-2}$. Thus in a sense Welch's result is true for the general linear model tested within the full model.

James (1951) considers the problem of testing n means equal to zero. In this case $P_0 = I$, $P_1 = 0$ and this gives $B = 0$ and $A = \Sigma_i 1/f_i$, which is consistent with James' result.

Notice how, in general, we can interprete the result in terms of the weighted regression with D^{-1} as weights. This gives rise to the projections P_0 and P_1 and P_{1ii}, say, is just the coefficient of Y_i in the expression for the fitted value in this regression $(P_1 Y)_{ii}$.

A different interpretation can be given as follows: Define the residuals in the regression with D^{-1} as weights by $R_1 = (I - P_1)Y$, $R_0 = (I - P_0)Y$ and $R_{01} = (P_0 - P_1)Y = R_1 - R_0$. Then as is well known $V(R_1) = (I - P_1)D$, $V(R_0) = (I - P_0)D$ and $V(R_{01}) = (P_0 - P_1)D$, which means that

(5) $$A = \Sigma_i \, V(R_{01i})V(R_{1i})/(V(Y_i))^2 f_i$$

(6) and $$B = \Sigma_i \, V(R_{01i})V(R_{0i})/(V(Y_i))^2 f_i$$

The results of Theorem 1 can be applied as Welch (1951) does to fit a distribution of the form $cF(f_1,f_2)$ such that it has the same mean and variance as Q.

This is accomplished if we take

(7) $$f_1 = p - m$$

(8) $$c = p - m + 2(A + B) - 6(A - B)/(p - m + 2)$$

(9) $$f_2 = (p - m)(p - m + 2)/3(A - B)$$

We have here used the approximations

$$EcF(f_1,f_2) = c(1 - \frac{2}{f_2})^{-1} = c(1 + \frac{2}{f_2})$$

$$VcF(f_1,f_2) = c^2(1 - \frac{2}{f_2})^{-1}(1 - \frac{4}{f_2})^{-1}(1 + \frac{2}{f_1}) - c^2(1 - \frac{2}{f_2})^{-2}$$

$$= c^2 \frac{2}{f_1}(1 + \frac{f_1+6}{f_2})$$

It is difficult to compare the results of (3) and (4) to the results of James (1954), due to the rather complicated form in which they are given in his paper.

The result can be formulated as follows

$$P\{Q \leqq \xi + h_1(\xi)\} = 1 - \alpha + O(\frac{1}{f_i^2})$$

where $\xi = \chi^2_{1-\alpha}(p-m)$, and $h_1(\xi)$ is of the order of $1/f_i$.

From the relation

$$F_{1-\alpha}(f_1,f_2) = \frac{\chi^2_{1-\alpha}(f_1)}{f_1} \quad (1 + \frac{1}{2f_2} \quad (\chi^2_{1-\alpha}(f_1) - f_1 + 2))$$

it follows that in our notation

$$h_1(\xi) = \frac{\xi}{2(p-m)}(A + 7B + 3(A - B)\xi/(p - m + 2)) \quad .$$

We shall go through a few special cases to show how easy $h_1(\xi)$ is found with the present notation.

Consider for instance a two way analysis of variance, where Y_{ij} has mean ξ_{ij} and variance σ_{ij}^2 $i=1,\ldots,r, j=1,\ldots,s$. We want to test the hypothesis of an additive model, i.e. $\xi_{ij}=\alpha_i+\beta_j$. In In order to find A and B we assume $w_{ij}=\sigma_{ij}^{-2}$ known and let $w_{i\cdot}=\Sigma_j w_{ij}, w_{\cdot j}=\Sigma_i w_{ij}$, and $w_{\cdot\cdot}=\Sigma_{ij}w_{ij}$, then the estimate for ξ_{ij} is given by

$$\hat{\xi}_{ij}=\Sigma_i Y_{ij}w_{ij}/w_{\cdot j}+\Sigma_j Y_{ij}w_{ij}/w_{i\cdot}-\Sigma_{ij}Y_{ij}w_{ij}/w_{\cdot\cdot}$$

which gives

$$P_{1\,ij,ij}=w_{ij}(1/w_{i\cdot}+1/w_{\cdot j}-1/w_{\cdot\cdot}) \quad .$$

Since $P_0=I$, we are testing inside the full model, we find $B=0$ and

$$A=\Sigma_{ij}(1-w_{ij}(1/w_{i\cdot}+1/w_{\cdot j}-1/w_{\cdot\cdot}))^2/f_{ij}$$

$$h_1(\xi)=\frac{\xi}{2(r-1)(s-1)}\left(1+\frac{3\xi}{(r-1)(s-1)+2}\right)$$

$$\Sigma_{ij}(1-w_{ij}(1/w_{i\cdot}+1/w_{\cdot j}-1/w_{\cdot\cdot}))^2/f_{ij}$$

This should be compared to (4.19) of James (1954).

Another case of interest is a simple linear regression where the mean of Y_i is $\alpha+\beta t_i$ and we let σ_i^2 denote the variance. We want to test $\beta=0$ and in order to find A and B we assume $w_i=\sigma_i^{-2}$ known. Then

$$(P_0Y)_i=\Sigma_j w_j Y_j/\Sigma_j w_j+(t_i-\bar{t})\Sigma_j w_j Y_j(t_j-\bar{t})/SSD$$

where $\bar{t} = \Sigma_j w_j t_j / \Sigma_j w_j$ and $SSD = \Sigma_j w_j (t_j - \bar{t})^2$

We also find

$$(P_1 Y)_i = \Sigma_j w_j Y_i / \Sigma_j w_j .$$

Hence

$$A = \Sigma_i w_i (t_i - \bar{t})^2 (1 - w_i/\Sigma_j w_j)/SSDf_i$$

$$B = \Sigma_i w_i (t_i - \bar{t})^2 (1 - w_i/\Sigma_j w_j - (t_i - \bar{t})^2 w_i/SSD)/SSDf_i$$

$$h_1(\xi) = \frac{\xi}{2}(\Sigma_i w_i (t_i - \bar{t})^2 (8(1 - w_i/\Sigma_j w_j) - 7(t_i - \bar{t})^2 w_i/SSD)/SSDf_i$$
$$+ \Sigma w_i^2 (t_i - \bar{t})^4/SSD^2 f_i)$$

As a byproduct of the above analysis we can also obtain various other approximations. Let $P(S)$ denote the projection onto a subspace L with respect to S^{-1} and let $P = P(D)$, then we have

<u>Theorem 2</u> Let $\hat{Y} = P(S)Y$ denote the fitted values, then $E\hat{Y} = \xi$ and up to terms of order $1/f_i$, we have

(10) $$V(\hat{Y})_{ij} = P_{ij}\sigma_j^2 + 2\Sigma_\nu P_{i\nu}(1 - P_{\nu\nu})P_{\nu j}\sigma_j^2/f_\nu$$

and

(11) $$E(P(S)S)_{ij} = P_{ij}\sigma_j^2 - 2\Sigma_\nu P_{i\nu}(1 - P_{\nu\nu})P_{\nu j}\sigma_j^2/f_\nu$$

<u>Corollary</u> An estimate of $V(Y)_{ij}$ which is unbiased up to terms of order $1/f_i$ is given by

(12) $$P_{ij}(S)s_j^2 + 4\Sigma_\nu P_{i\nu}(S)(1 - P_{\nu\nu}(S))P_{\nu j}(S)s_j^2/f_\nu$$

These results generalize the formula for the variance in a simple regression model, as derived by Jacquez, Mather and Crawford (1968).

Finally we shall note the following: The expressions in Theorem 1 and 2 all depend on the true but unknown D. If we insert S in place of D, that is replace P_ν by $P_\nu(S)$, then a bias of order $1/f_i^2$ will result in A and B. With the accuracy we are working with this does not matter.

3. Proof of main result

We shall first prove the results for $D = I$ and then transform to the general case. Welch (1951) derived the first terms of the moment generating function but we shall only need the first two moments to get the desired accuracy.

Thus we want to find first the mean and variance of Q given S.

(13) $$E(Q|S) = E(Y'TY|S) = \mathrm{tr}T$$

(14) $$V(Q|S) = V(Y'TY|S) = 2\mathrm{tr}T^2$$

We shall need the following approximation result.

Lemma Let P(S) be a projection on to the subspace L with respect to S^{-1}, then if $S^{-1} = I + U$, we have

$$P(S) = P + PU(I-P) - PUPU(I-P) + \varepsilon$$

where $\varepsilon \in o\,(\|U\|^2)$

Proof Let $\xi \in L \Leftrightarrow \xi = X\beta, \beta \in R^p$, and rank $X = p$, then

$$\begin{aligned} P(S) &= X(X'S^{-1}X)^{-1}X'S^{-1} \\ &= X(X'X + X'UX)^{-1}X'(I+U) \\ &= X((X'X)^{-1} - (X'X)^{-1}X'UX(X'X)^{-1} \\ &\quad + (X'X)^{-1}(X'UX(X'X)^{-1})^2 + \varepsilon)X'(I+U) \end{aligned}$$

from which the result follows for $P = X'(X'X)^{-1}X$.

Now we apply this result to $T = S^{-1}(P_o(S) - P_1(S))$. This gives

$$T = P_o - P_1 + (P_o - P_1)U(I - P_o) + (I - P_1)U(P_o - P_1)$$
$$- (I - P_1)UP_1U(I - P_1) + (I - P_o)UP_oU(I - P_o) + \varepsilon$$

and hence

(15) $$\mathrm{tr}T = p - m + 0 + \Sigma_i U_{ii}(P_{oii} - P_{1ii}) - \Sigma_{ij} U_i U_j P_{1ij}(\delta_{ji} - P_{1ji})$$
$$+ \Sigma_{ij} U_i U_j P_{oij}(\delta_{ji} - P_{oji}) + \varepsilon$$

Now we can find the expectation of Q as $E(EQ|S) = E\,\mathrm{tr}T$. Note therefore that

$$EU_i = E(S_i^{-1} - 1) = 1/(1 - 2/f_i) - 1 = 2/f_i$$

$$EU_i^2 = E(S_i^{-1} - 1)^2 = 1/[(1 - \frac{2}{f_i})(1 - \frac{4}{f_i})] - 2/(1 - \frac{2}{f_i}) + 1 = 2/f_i$$

and that EU_iU_j is of order $1/(f_if_j)$ and therefore discarded if $i \neq j$.

Then

$$EQ = p - m + \Sigma_i 2(P_{oii} - P_{1ii} - P_{1ii}(1 - P_{1ii}) + P_{oii}(1 - P_{oii}))/f_i$$
$$= p - m + 2(A + B).$$

Similarly we find

(16) $$\mathrm{tr}T^2 = \mathrm{tr}(P_o - P_1) + 2\mathrm{tr}U(P_o - P_1) + \mathrm{tr}U(P_o - P_1)U(P_o - P_1)$$
$$- 2\mathrm{tr}\; UP_1U(P_o - P_1) + 2\mathrm{tr}\; U(I - P_o)U(P_o - P_1)$$

and hence

$$EV(Q|S) = 2[p - m + \Sigma_i 2(P_{oii} - P_{lii})(2 + P_{oii} - P_{lii} - 2P_{lii} + 2(1 - P_{oii}))/f_i]$$

From (15) we also find

$$V(E(Q|S)) = V(trT) = \Sigma_i 2(P_{oii} - P_{lii})^2/f_i$$

and adding these we end up with

$$\begin{aligned} V(Q) &= EV(Q|S) + V(EQ|S) \\ &= 2(p - m) + \Sigma_i (P_{oii} - P_{lii})(16 - 2P_{oii} - 14P_{lii})/f_i \\ &= 2(p - m) + 14A + 2B \end{aligned}$$

If $D \neq I$, consider the variable $\tilde{Y} = D^{-\frac{1}{2}}Y$ with mean $D^{-\frac{1}{2}}\xi$ and variance matrix I. Then let $\tilde{S} = D^{-\frac{1}{2}}SD^{-\frac{1}{2}}$. The projection $\tilde{P}(\tilde{S}) = D^{-\frac{1}{2}}P(S)D^{\frac{1}{2}}$ and $\tilde{P} = D^{-\frac{1}{2}}PD^{\frac{1}{2}}$ or $P_{ij} = \sigma_i^{-1}P_{ij}\sigma_j$ and in particular the diagonal elements are not changed. It is seen that as well Q as A and B are invariant under this transformation.

To prove Theorem 2 on V(P(S)Y) note that

$$\begin{aligned} V(P(S)Y) &= EV(P(S)Y|S) + V(E(P(S)Y|S) \\ &= EP(S)P(S)' + 0 \end{aligned}$$

which by the Lemma can be written

$$E(P + PU(I - P)UP + \ldots)$$

and hence

$$V(P(S)Y)_{ij} = P_{ij} + 2\Sigma_\nu P_{i\nu}(1 - P_{\nu\nu})P_{\nu j}/f_\nu$$

which shows (10) for the case $D = I$.

Finally the Lemma applied to $P(S)S$ gives

$$EP(S)S = E(P - PUP + PUPUP + \ldots)$$

and hence

$$EP_{ij}(S)s_j^2 = P_{ij} - 2\Sigma_\nu \; P_{i\nu}(1 - P_{\nu\nu})P_{\nu j}/f_\nu$$

which shows (11). The general case $D \neq I$ is solved as before by applying (10) and (11) to the variables $\tilde{Y}$ and $\tilde{S}$. Finally the corollary follows easily from Theorem 2.

4. Generalizations

The methods above give a much more general result. Let us first define $Z = S - I$, then $U = S^{-1} - I = (I + Z)^{-1} - I = - Z + Z^2 + \ldots$. We can now collect the results of (13) - (16) and we find that to the order of approximation we need:

(17) $$EQ = p - m + E\text{tr}\{Z(P_o - P_1)Z(P_o - P_1)\} + 2E\text{tr}\{Z(P_o - P_1)Z(I - P_o)\}$$

(18) $$VQ = 2(p - m) + E\{\text{tr}(Z(P_o - P_1))\}^2 + 6E\;\text{tr}\{Z(P_o - P_o)Z(P_o - P_o)\} + 8E\;\text{tr}\{Z(P_o - P_1)Z(I - P_o)\}$$

This result does not depend on any particular form of the estimate S, only on the fact that certain moments exist.

We shall consider two generalizations. First let $V_1, \ldots, V_K$ denote an orthogonal decomposition of R^n with dimensions $m_1, \ldots, m_K$ and projections $Q_1, \ldots, Q_K$. Consider the family of distributions of Y on R^n given by the normal distribution with mean $\xi \in L_o$ a subspace of R^n and variance $\Gamma = \Sigma_{i=1}^K \; \Gamma_i$, where Γ_i is an arbitrary covariance matrix giving rise to a normal distribution with support on V_i. Note that this implies that $Q_i \Gamma_i Q_j = \delta_{ij} \Gamma_i$. We want to test the hypothesis that $\xi \in L_1 \subset L_o$.

Further we shall let S_i denote an estimate of Γ_i which has a Wishart distribution of the form $W_{m_i}(f_i, \Gamma_i)/f_i$. In this situation we can complete the reduction of the expressions (17) and (18) by means of the following Lemma.

Lemma Let S have the distribution $W_p(f, \Sigma)/f$ and let M and N denote symmetric $p \times p$ matrices then

$$E\ tr\{(S-\Sigma)M(S-\Sigma)N\} = E\ tr(S-\Sigma)M\ tr(S-\Sigma)N$$

$$= \{tr\ M\Sigma N\Sigma + (tr\ M\Sigma)(tr\ N\Sigma)\}/f$$

Proof Let $U_1, \ldots, U_f$ be independent normally distributed with mean 0 and covariance matrix Σ, then $S = \Sigma_1^f U_i U_i'/f$ and

$$E\ tr(S-\Sigma)M(S-\Sigma)N$$

$$= \Sigma_i \Sigma_j\ E\ tr(U_i U_i' - \Sigma)M(U_j U_j' - \Sigma)N/f^2$$

$$= E\ tr(U_1 U_1' - \Sigma)M(U_1 U_1' - \Sigma)N/f$$

$$= (E\ U_1' M U_1 U_1' N U_1 - tr\ \Sigma M \Sigma N/f$$

$$= (tr\ \Sigma M \Sigma N + tr\ \Sigma M\ tr\ \Sigma N)/f$$

Similarly

$$E\ tr(S-\Sigma)M\ tr(S-\Sigma)N$$

$$= \Sigma_i \Sigma_j\ E\ tr(U_i U_i' - \Sigma)M\ tr(U_j U_j' - \Sigma)N/f^2$$

$$= E\ \{tr(U_1 U_1' - \Sigma)M\ tr(U_1 U_1' - \Sigma)N\}/f$$

$$= \{E\ U_1' M U_1 U_1' N U_1 - tr\ \Sigma M \Sigma N\}/f$$

as was to be proven.

From this Lemma it follows that if $S = \sum_{i=1}^{K} S_i$ then $V = S - I = \sum_{i=1}^{K} (S_i - Q_i)$ and it follows that the result of Theorem 1 holds with

(19) $A = \frac{1}{2} \Sigma_i \{tr(P_o - P_1)Q_i(I - P_1)Q_i + tr(P_o - P_1)Q_i\, tr(I - P_1)Q_i\}/f_i$

(20) $B = \frac{1}{2} \Sigma_i \{tr(P_o - P_1)Q_i(I - P_o)Q_i + tr(P_o - P_1)Q_i\, tr(I - P_o)Q_i\}/f_i$

As a simple example of this consider the situation where $Y_1,\ldots,Y_K$ are independent p-dimensional normally distributed with means $\xi_1,\ldots,\xi_K$ and covariance matrices $\Gamma_1,\ldots,\Gamma_K$. We want to test the hypothesis that $\xi_1 = \ldots = \xi_K$.

The estimate of the common mean is $\hat{\xi} = (\Sigma_i \Gamma_i^{-1})^{-1} \Sigma_i \Gamma_i^{-1} \xi_i$. To put it into the above framework let $R^{Kp} = R^P \times \ldots \times R^P$ and let Q_i denote the projection into the i'th copy of R^P considered as a subspace of R^{Kp}. The projection $P(\Gamma)$ can be represented as a block matrix with (s,t)'th block equal to $W^{-1}W_t$ $(W_t = \Gamma_t^{-1},\ W = \Sigma_t W_t)$, $s,t = (1,\ldots,k)$. Similarly $(Q_i)_{s,t} = I_{p\times p}$ if $s = t = i$ and zero otherwise. Hence we find for $P_o = I$, that $B = 0$ and

$$A = \tfrac{1}{2} \Sigma_i [tr\{(I - P(\Gamma))Q_i\}^2 + \{tr(I - P(\Gamma))Q_i\}^2]/f_i$$

$$= \tfrac{1}{2} \Sigma_i\ tr(I - W^{-1}W_i)^2 + \{tr(I - W^{-1}W_i)\}^2]/f_i$$

which is consistent with the result of James (1954).

As the final generalization we shall consider the case where $\Gamma_i = \sigma_i^2 Q_i$, and where σ_i^2 is estimated by s_i^2 which has a χ^2 distribution with f_i degrees of freedom and a scale parameter σ_i^2/f_i. Also let $s_1^2,\ldots,s_k^2$ be independent. Then we find $V = S - I = \Sigma_i (s_i^2 - 1)Q_i$ and

$$E\ tr\{(S-I)M(S-I)N\} = \Sigma_{ij}\ tr(Q_iMQ_jN)E(s_i^2-1)(s_j^2-1)$$

$$= \Sigma_i\ 2tr(Q_iMQ_iN)/f_i$$

and

$$E\ tr\{(S-I)M\ tr(S-I)N\} = \Sigma_{ij}\ tr(Q_iM)tr(Q_jN)E(s_i^2-1)(s_j^2-1)$$

$$= \Sigma_i\ 2tr(Q_iM)tr(Q_iN)/f_i$$

Hence we find that

$$EQ = p - m + 2A + 2B$$

$$VQ = 2(p-m) + 14A + 2B + 2C$$

where

$$A = \Sigma_i\{tr\ Q_i(P_o - P_1)Q_i(I - P_1)\}/f_i$$

$$B = \Sigma_i\{tr\ Q_i(P_o - P_1)Q_i(I - P_o)\}/f_i$$

$$C = \Sigma_i[\{tr\ (Q_i(P_o - P_1))\}^2 - tr\{Q_i(P_o - P_1)\}^2]/f_i$$

As an application of this result consider the balanced incomplete block design given by

		Treatment		
		A	B	C
	1	Y_1	Y_2	
Block	2	Y_3		Y_4
	3		Y_5	Y_6

There are three subspaces of the 6-dimensional observations space, which are of interest. These spaces are spanned by vectors of the form:

Block space

α	α	
β		β
	γ	γ

Plot space

α	$-\alpha$	
β		$-\beta$
	γ	$-\gamma$

Treatment space

α	β	
α		γ
	β	γ

The projections onto the block space and plot space Q_b and Q_p are best described as block diagonal matrices with 2×2 matrices along the diagonal. These are $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ and $\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$ respectively.

Thus the covariance matrix is

$$\Gamma = \tau^2 Q_b + \sigma^2 Q_p, \quad \sigma^2 > 0, \quad \tau^2 > 0,$$

and the projection with respect to $\Gamma^{-1} = \tau^{-2} Q_b + \sigma^{-2} Q_p$ onto the treatment space L_0 is found to be

$$\left[\begin{array}{cc|cc|cc} \frac{1}{2} & \phi & \frac{1}{2} & \phi & -\phi & -\phi \\ \phi & \frac{1}{2} & -\phi & -\phi & \frac{1}{2} & \phi \\ \hline \frac{1}{2} & \phi & \frac{1}{2} & \phi & -\phi & -\phi \\ -\phi & -\phi & \phi & \frac{1}{2} & \phi & \frac{1}{2} \\ \hline \phi & \frac{1}{2} & -\phi & -\phi & \frac{1}{2} & \phi \\ -\phi & -\phi & \phi & \frac{1}{2} & \phi & \frac{1}{2} \end{array}\right], \quad \phi = (\sigma^2 - \sigma^2)/2(3\tau^2 + \sigma^2)$$

If we want to test the hypothesis of no treatment effects, we also have to project onto the space L_1 spanned by the vector of 1's. Since L_1 is contained in the blockspace the projection can be written $(P_1 Y)_i = \bar{Y}$. Now let s_1^2 and s_2^2 be estimates of τ^2 and σ^2 with f_1 and f_2 degrees of freedom. Then we can test the hypothesis of no treatment effect using the information from both strata by means of the statistic Q. Some calculations show that in this case the corrections to the mean and variance are given by

$$A = (2\sigma^2/f_1 + 6\tau^2/f_2)/(3\tau^2 + \sigma^2)$$

$$B = (1/f_1 + 1/f_2)\,6\tau^2\sigma^2/(3\tau^2 + \sigma^2)^2$$

$$C = 2\{\sigma^4/f_1 + 9\tau^4/f_2\}/(3\tau^2 + \sigma^2)^2$$

5. Acknowledgement

It is a pleasure to thank Anders Hald for useful comments to a first draft of the paper.

References

(1) James, G.S. (1951). The comparison of several groups of observations when the ratioes of the population variances are unknown. Biometrika 38,324-329.

(2) James, G.S. (1954). Tests of linear hypothesis in univariate and multivariate analysis when the ratioes of the population variances are unknown. Biometrika 41, 19-43.

(3) Welch, B.L. (1947). The generalisation of "Students" problem when several different population variances are involved. Biometrika 34,28-35.

(4) Welch, B.L. (1951). On the comparison of several mean values: an alternative approach. Biometrika 38, 330-336.

(5) Jacquez, J.A., Mather, F.J., and Crawford, C.R. (1968). Linear regression with non-constant, unknown error variances: Sampling experiments with least squares and maximum likelihood estimators. Biometrics **24**, 607-626.

PREPRINTS 1978

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Tjur, Tue: Statistical Inference under the Likelihood Principle.

No. 2 Hering, Heinrich: The Non-Degenerate Limit for Supercritical Branching Diffusions.

No. 3 Henningsen, Inge: Estimation in M/G/1-Queues.

No. 4 Braun, Henry: Stochastic Stable Population Theory in Continuous Time.

No. 5 Asmussen, Søren: On some two-sex population models.

No. 6 Andersen, Per Kragh: Filtered Renewal Processes with a Two-Sided Impact Function.

No. 7 Johansen, Søren & Ramsey, Fred L.: A Bang-Bang Representation for 3x3 Embeddable Stochastic Matrix.

No. 8 Braun, Henry: A Simple Method for Testing Goodness of Fit in the Presence of Nuisance Parameters.

No. 9 Lauritzen, S.L. & Speed, T.P. & Vijayan, K.: Decomposable Graphs and Hypergraphs.

No. 10 Hald, Anders: On the Statistical Theory of Sampling Inspection by Attributes.

No. 11 Darroch, J.N. & Lauritzen, S.L. & Speed, T. P.: Markov Fields and Log-linear Interaction Models for Contingency Tables.

No. 12 Speed, T.P.: A Note on Nearest-Neighbour Gibbs and Markov Probabilities.

No. 13 Andersson, Steen A.: Canonical Correlations with Respect to a Complex Structure.

PREPRINTS 1979

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Edwards, David: Large Sample Tests for Stationarity and Reversibility in Finite Markov Chains.

No. 2 Andersson, Steen A. : Distribution of Maximal Invariants Using Proper Action and Quotient Measures.

No. 3 Johansen, Søren: A Note on the Welch-James Approximation to the Distribution of the Residual Sum of Squares in a Weighted Linear Regression.