Tema 4. Diseño por bloques aleatorizado

1. Introducción

El análisis por bloques permite controlar sistemáticamente fuentes de variabilidad externas que intervienen o actuan en la observación de la variable respuesta. Puesto que en estos casos, la componente aleatoria de error o error experimental asociada a cada observación refleja tanto el error aleatorio como la variabilidad debida a los elementos físicos que intervienen en la realización del experimento. Por tanto, se considerará no sólo el diseño de factores internos sino también la actuación de las fuentes de variabilidad externas que se reflejan en el diseño de los bloques o variación entre bloques.

2. Diseño por bloques aleatorizados completos

El modelo estadístico que se considera en este diseño viene dado por

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \quad i = 1, \ldots, k, \; j = 1, 2, \ldots, b, \tag{1}$$

siendo μ la media general, τ_i representa el efecto del $i-$ésimo tratamiento, β_j es el efecto del $j-$ésimo bloque y ϵ_{ij}, para $i = 1, \ldots, k$, y $j = 1, \ldots, b$, son las componentes de error aleatorio que se suponen normalmente distribuidas e independientes. Consideraremos primeramente el caso de efectos fijos, es decir, que el efecto de los tratamientos recogidos en los parámetros τ_i para $i = 1, \ldots, k$, y de los bloques recogidos en los parámetros β_j para $j = 1, 2, \ldots, b$, son fijos, interpretándose como desviaciones respecto a la media, por tanto,

$$\sum_{i=1}^{k} \tau_i = 0, \quad \sum_{j=1}^{b} \beta_j = 0.$$

En primer lugar, se considera el contraste sobre la igualdad de medias de los tratamientos que es equivalente a contrastar:

$$\begin{aligned} H_0 &: \tau_1 = \tau_2 = \cdots = \tau_k = 0 \\ H_1 &: \tau_i \neq 0, \quad \text{al menos para algún i.} \end{aligned}$$

A partir de las siguientes cantidades

$$\begin{aligned} y_{i\cdot} &= \sum_{j=1}^{b} y_{ij}, \quad i = 1, \ldots, k \\ y_{\cdot j} &= \sum_{i=1}^{k} y_{ij}, \quad j = 1, \ldots, b \\ y_{\cdot\cdot} &= y_{ij} = \sum_{j=1}^{b} y_{\cdot j} = \sum_{i=1}^{k} y_{i\cdot}, \\ \bar{y}_{\cdot\cdot} &= \frac{1}{kb} y_{\cdot\cdot} \end{aligned} \tag{2}$$

tras operar en la cantidad

$$\sum_{i=1}^{k} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{\cdot\cdot})^2$$

se obtiene

$$\begin{aligned} \sum_{i=1}^{k} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{\cdot\cdot})^2 &= b \sum_{i=1}^{k} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 \\ &+ k \sum_{j=1}^{b} (\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2 + \sum_{i=1}^{k} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{\cdot j} - \bar{y}_{i\cdot} + \bar{y}_{\cdot\cdot})^2 \end{aligned}$$

la descomposición de la variabilidad total en la suma de la variabilidad entre tratamientos, variabilidad entre bloques y variabilidad debida al error. Dado que se parte de N observaciones, se tienen $N - 1$ grados de libertad para la suma total de cuadrados y $k - 1$ y $b - 1$ grados de libertad para la suma de cuadrados para los tratamientos y para los bloques, puesto que se han considerado k tratamientos y b bloques. La variabilidad o suma de cuadrados debida al error se obtiene restando a la suma global, la suma sobre tratamientos y la suma sobre bloques, es decir,

$$SS_{TOTAL} = SS_{TRATAMIENTOS} + SS_{BLOQUES} + SS_{ERROR},$$

por tanto,

$$SS_{ERROR} = SS_{TOTAL} - SS_{TRATAMIENTOS} - SS_{BLOQUES}.$$

Los grados de libertad de SS_{ERROR} son $(kb-1)-(k-1)-(b-1)$. A partir de la suposición de normalidad e independencia de los errores, se obtiene que $SS_{TRATAMIENTOS}/\sigma^2$, $SS_{BLOQUES}/\sigma^2$ y SS_{ERROR}/σ^2 son variables aleatorias independientes con distribuciones χ^2. Adicionalmente, el valor medio o esperanza de las sumas de cuadrados divididas por sus grados de libertad (medias de cuadrados) es el siguiente

$$\begin{aligned} E[MS_{TRATAMIENTOS}] &= \sigma^2 + \frac{b\sum_{i=1}^{k}\tau_i^2}{k-1} \\ E[MS_{BLOQUES}] &= \sigma^2 + \frac{k\sum_{j=1}^{b}\beta_j^2}{b-1} \\ E[MS_{ERROR}] &= \sigma^2. \end{aligned}$$

Para el contraste de igualdad de medias de los tratamientos se construye entonces el estadístico

$$F_0 = \frac{MS_{TRATAMIENTOS}}{MS_{ERROR}},$$

que se distribuye según una F con $k-1$ y $(kb-k-b+1)$ grados de libertad. Se tiene entonces que se rechaza la igualdad de medias de los tratamientos cuando $F_0 > F_{\alpha,kb-k-b+1}$ para $F_{\alpha,kb-k-b+1}$ el percentil $1-\alpha$ de la distribución F con $k-1$ y $(kb-k-b+1)$ grados de libertad.

Otro aspecto a analizar es si existen diferencias importantes entre las medias por bloques, puesto que en caso contrario, el diseño por bloques puede que no sea necesario, al no proporcionar en el análisis por bloques una reducción importante del factor de ruido, que se reflejaría en un aumento de la precisión de la comparación entre las medias de los tratamientos. La dificultad en el diseño de contrastes sobre la igualdad de medias por bloques reside en que la aleatorización se aplica sólo a los tratramientos dentro de cada bloque y, por tanto, la determinación de bloques representa una restricción para la aleatorización. Existen diferentes interpretaciones en relación con la aplicabilidad o validez del estadístico

$$F_0 = \frac{MS_{BLOQUES}}{MS_{ERROR}}$$

para el contraste de medias entre bloques. Más concretamente, se considera que la aplicabilidad de este estadístico está relacionada con la hipótesis de aleatorización. En este sentido, la restricción del diseño por bloques resta validez a este estadístico. Por otra parte, cuando los errores se distribuyen según normales independientes, este estadístico puede ser considerado para comparar las medias por bloques. En general, se concluye que $F_0 = \frac{MS_{BLOQUES}}{MS_{ERROR}}$ proporciona un procedimiento aproximado para estudiar e efecto de la variable bloque.

3. Otros aspectos del análisis estadístico del diseño por bloques

3.1. Carácter aditivo del modelo.

El carácter aditivo del modelo en relación con el efecto de los tratamientos y bloques no siempre es adecuado. En ocasiones, se detecta una relación multiplicativo en la escala original

$$E[y_{ij}] = \mu\tau_i\beta_j,$$

que se transforma en una relación aditiva o lineal en la escala logarítmica, por ejemplo,

$$\ln E[y_{ij}] = \ln\mu + \ln\tau_i + \ln\beta_j,$$

o equivalentemente,

$$E[y_{ij}^*] = \mu^* + \tau_i^* + \ln\beta_j^*.$$

No siempre se puede eliminar mediante una transformación adecuada el esquema no aditivo (interacción bloque-tratamiento). Este hecho se puede detectar mediante un análisis de los residuos junto con otros procedimentos de diagnosis. En ocasiones, cuando un análisis de la interacción entre bloques y tratamientos es necesario, se puede considerar un diseño factorial que será objeto de estudio en temas posteriores.

3.2. Modelo de efectos aleatorios.

En relación con el modelo de efectos fijos (1), se puede considerar su versión aleatoria considerando que los tratamientos y los bloques, o bien, alguno de ellos son variables aleatorias. El análisis estadístico es similar, aunque se aprecian cambios importantes en la interpretación. Más concretamente, si los bloques son aleatorios es esperable que la comparación de tratamientos sea la misma en toda la población de bloques de la que se han extraído los bloques muestrales que intervienen en la realización de los experimentos. La esperanza de la media de cuadrados de los bloques viene dada, en este caso, por la siguiente expresión:

$$E[MS_{BLOQUES}] = \sigma^2 + k\sigma_\beta^2,$$

siendo σ_β^2 la componente de varianza del efecto de los bloques. En cualquier caso, la media de cuadrados sobre los tratamientos siempre es independiente del efecto de los bloques, considerándose el estadístico

$$F_0 = MS_{TRATAMIENTOS}/MS_{ERROR}$$

para el contraste de variabilidad entre los tratamientos. En el caso en el que existen interacciones significativas entre los tratamientos y bloques, la prueba de comparación de medias entre tratamientos no se ve afectada cuando los bloques son aleatorios, dado que la media de cuadrados de los tratamientos y del error contienen el efecto de la interacción. Así, se mantiene también, en este caso de interacción, el estadístico para la comparación de las medias de los tratamientos, puesto que dicho estadístico no refleja el efecto de la interacción.

3.3. Comparaciones múltiples

En el caso de modelos de efectos fijos, donde los tratamientos y bloques no son aleatorios, cuando las medias de los tratamientos presentan diferencias significativas, suele ser necesario identificar cuáles son los tatamientos que presentan niveles medios diferentes. Para el planteamiento de contrastes múltiples, cualquiera de los tests planteados en el análisis de la varianza de una sóla vía son válidos. La metodología ANOVA asociada a un sólo factor se adapta a esta situación sustituyendo el parámetro número de réplicas n por el número de bloques b. Se deben corregir asimismo los grados de libertad del error de la metodología ANOVA de una sóla vía, sustituyendo $k(n-1)$ por $(k-1)(b-1)$.

3.4. Elección del tamaño muestral

La determinación del tamaño muestral o número de réplicas es muy importante en el diseño aleatorizado por bloques, dado que el número de bloques determina el número de réplicas o tamaño muestral b, que al aumentar afecta a los grados de libertad de la media de cuadrados del error. La metodología ANOVA de una sóla vía para la selección del número de réplicas en la realización de un experimento completamente aleatorizado de un factor puede ser adaptada al contexto del diseño aleatorizado por bloques en relación con el número de bloques para cada tratamiento. Así las curvas de operación característica ANOVA de una sóla vía pueden reformularse como sigue

$$\Phi^2 = \frac{b\sum_{i=1}^{k}\tau_i^2}{k\sigma^2}.$$

En el caso de factor o tratamiento aleatorio se tiene

$$\lambda = \sqrt{1+\frac{b\sigma_\tau^2}{\sigma^2}}.$$

3.5. Eficiencia del diseño aleatorizado por bloques

La comparación del diseño aleatorizado por bloques en relación con el diseño completamente aleatorizado se suele hacer en términos de eficiencia relativa mediante la fórmula

$$ER = \frac{(df_b+1)(df_{CA}+3)\sigma_{CA}^2}{(df_b+3)(df_{CA}+1)\sigma_b^2},$$

siendo σ_{CA}^2 y σ_b^2 las varianzas del error experimental del diseño completamente aleatorizado y del diseño completamente aleatorizado por bloques y df_b y df_{CA} los correspndientes grados de libertad del error. Esta ecuación puede interpretarse como el incremento de réplicas necesario para que un diseño completamente aleatorizado pueda ser utilizado en lugar de un diseño aleatorizado por bloques, manteniendo la misma sensibilidad en ambos diseños. Se realiza asimismo un ajuste sobre la diferencia de grados de libertad entre los dos diseños mediante la razón de grados de libertad en ER. Una estimación de la varianza del error en el diseño completamente aleatorizado

σ^2_{CA} en términos de la varianza de la suma de cuadrados medios del error en el diseño aleatorizado por bloques viene dada por

$$\widehat{\sigma}^2_{CA} = \frac{(b-1)MS_{BLOQUES} + b(k-1)MS_{ERROR}}{ab-1}.$$

4. Adecuación del modelo

En la determinación de la idoneidad del modelo, los aspectos que deben ser estudiados son los siguientes:

- Hipótesis de normalidad
- Desigualdad entre las varianzas del error en los bloques o tratamientos
- Interacción entre bloques y tratamientos.

Para estudiar estos tres aspectos la herramienta fundamental es el análisis de los residuos, que vienen dados por

$$e_{ij} = y_{ij} - \widehat{y}_{ij},$$

siendo

$$\widehat{y}_{ij} = y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot}.$$

En la revisión de la hipótesis de normalidad se suelen utilizar métodos gráficos, así como pruebas de ajuste. En relación con la desigualdad de varianzas se suele graficar los residuos y comprobar si presentan diferentes niveles de dispersión o variabilidad. Finalmente, para determinar si existen interacciones significativas entre los tratamientos y los bloques, se suele realizar una gráfica de los residuos frente a los valores ajustados $\widehat{y}_{ij}$. En principio, no debe existir relación entre la magnitud de los residuos y los valores ajustados, cuando no existen interacciones entre tratamientos y bloques.

4.1. Datos omitidos

Esta situación se presenta cuando se produce la pérdida de algún dato ocasionada por la no observación de la respuesta debida a la rotura de algún material, situación fuera de control, o bien, debido a alguna otra causa. Para abordar esta situación se puede sustiyuir el dato faltante por su estimación y proceder con el análisis de la varianza usual considerando el dato estimado como real y reduciendo los grados de libertad del error en una unidad, obteniéndose un análisis aproximado.

La estimación x del dato faltante se realiza a partir de la minimización de su contribución a la suma de cuadrados del error. Es decir x se define como el valor que minimiza la expresión

$$SS_{ERROR} = x^2 - \frac{1}{b}(y'_{i\cdot} + x)^2 - \frac{1}{k}(y'_{\cdot j} + x)^2 + \frac{1}{kb}(y'_{\cdot\cdot} + x^2) + R,$$

donde $'$ denota la suma con un dato faltante y R incluye todos los términos que no contienen a x. Se obtiene entonces como solución

$$x = \frac{ay'_{i\cdot} + by'_{\cdot j} - y'_{\cdot\cdot}}{(a-1)(b-1)}$$

como un estimador para la observación faltante. La ecuación anterior puede utilizarse iterativamente para estimar dos o más datos omitidos.

5. Estimación mínimo-cuadrática de modelo y contrastes de significación

Se considera nuevamente el modelo lineal

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \quad i = 1, \ldots, k, \; j = 1, 2, \ldots, b.$$

Bajo las restricciones

$$\sum_{i=1}^{k} \hat{\tau}_i = 0, \quad \sum_{j=1}^{b} \hat{\beta}_j = 0,$$

las soluciones del sistema de ecuaciones normales asociadas a la estimación mínimo-cuadrática de dicho modelo se definen como sigue:

$$\begin{aligned} \widehat{\mu} &= \bar{y}_{\cdot\cdot} \\ \widehat{\tau}_i &= \bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot} \quad i = 1, \ldots, k \\ \widehat{\beta}_j &= \bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot} \quad j = 1, \ldots, b. \end{aligned}$$

Sustituyendo los estimadores mínimo-cuadráticos en las ecuaciones normales se obtienen las estimaciones de los valores de la variable respuesta que vienen dados por

$$\widehat{y}_{ij} = \bar{y}_{i\cdot} + \bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot}.$$

En el planteamiento de tests de significación de los efectos de los tratamientos y de los bloques se procederá como en la regresión lineal múltiple. Más concretamente, para contrastar, por ejemplo, $H_0: \ \tau_i = 0$, se calcula la variabilidad explicada o reducción de la suma de cuadrados debida al ajuste del modelo

$$y_{ij} = \mu + \beta_j + \epsilon_{ij},$$

que viene dada por

$$R(\mu, \beta) = \sum_{j=1}^{b} \frac{y_{\cdot}^2 j}{k}$$

con b grados de libertad. Por tanto, la suma de cuadrados debida a $\{\tau_i\}$, tras introducir los parámetros μ y $\{\beta_j\}$ en el modelo es

$$R(\tau/\mu, \beta) = R(\mu, \tau, \beta) - R(\mu, \beta) = \sum_{i=1}^{k} \frac{y_{i\cdot}^2}{b} - \frac{y_{\cdot\cdot}^2}{kb}.$$

El estadístico para contrastar $H_0: \ \tau_i = 0$ vendría entonces definido mediante

$$F_0 = \frac{R(\tau/\mu, \beta)/(k-1)}{\left[\sum_{i=1}^{k} \sum_{j=1}^{b} y_{ij}^2 - R(\mu, \tau, \beta)\right]/(N-k)}.$$

Similarmente se procedería en el contraste de significación para los bloques $H_0: \ \beta_j = 0$, para $j = 1, \ldots, b$.

Nota. Para el problema de datos faltantes se ha planteado un procedimiento aproximado que introduce un sesgo en la media de cuadrados de los tratamientos que conduce a un exceso en el número de análisis que resultan significativos. Para un análisis exacto del problema de datos omitidos, que compense el desequilibrio producido por el hecho de que los bloques y tratamientos ya no son ortogonales (no hay el mismo número de tratamientos en todos los bloques), se puede utilizar el contraste de significación de la regresión general.

6. Diseño aleatorizado por bloques incompletos

En ocasiones no pueden realizarse todos los ensayos que corresponden a todas las posibles combinaciones de tratamientos dentro de cada bloque. Este hecho se suele dar debido a la escasez de medios, o bien, al tamaño físico de los bloques. Una solución es considerar un diseño aleatorizado por bloques en el que cada tratamiento no está presente en todos los bloques. Este tipo de diseño se conoce como diseño aleatorizado por bloques incompletos.

6.1. Diseño aleatorizado por bloques incompletos balanceado

Este tipo de diseño aleatorizado por bloques incompletos se establece bajo la condición de que cualquier par de tratamientos ocurran juntos el mismo número de veces. Se tiene entonces que si hay una restricción de k tratamientos por bloque, con $k < a$, un diseño balanceado por bloques incompletos se puede construir considerando $\begin{pmatrix} a \\ k \end{pmatrix}$ bloques y asignando una combinación diferente de tratamientos a cada bloque.

Para el análisis estadístico de este diseño se consideran a tratamientos y b bloques, con k tratamientos probados por bloque, repitiéndose r veces cada tratamiento en el diseño. Por tanto, el número total de observaciones es $N = ar = bk$. El número de bloques en los que un par de tratamientos ocurren juntos. es

$$\lambda = \frac{r(k-1)}{a-1},$$

siendo el diseño simétrico si $a = b$.

Ejemplo

Para $a = 4$ tratamientos, $b = 4$ bloques, $k = 2$ tratamientos por bloque, se puede construir un diseño por bloques incompleto balanceado de la siguiente forma

Bloques	Tratamientos
1	1,2,3
2	1,2,4
3	1,3,4
4	2,3,4

En este caso $\lambda = 2$.

Bajo el escenario paramétrico anterior se considera entonces el modelo

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}.$$

Dado que cada tratamiento ocurre en un conjunto diferente de r bloques, es necesaria la corrección de la suma de cuadrados de los tratamientos para separar los efectos de los tratamientos y el bloque. Dicha suma de cuadrados sobre tratamientos corregida viene dada por

$$SS_{TRATAMIENTOS\ CORREGIDA} = \frac{k\sum_{i=1}^{a} Q_i^2}{\lambda a},$$

donde Q_i es el total corregido del $i-$ésimo tratamiento, que viene dado mediante la fórmula:

$$Q_i = y_{i\cdot} - \frac{1}{k}\sum_{j=1}^{b} n_{ij} y_{\cdot j}, \quad i = 1, \ldots, a,$$

siendo $n_{ij} = 1$ si el tratamiento i ocurre en el bloque j y $n_{ij} = 0$ en caso contrario, es decir,

$$\frac{1}{k}\sum_{j=1}^{b} n_{ij} y_{\cdot j}$$

representa el promedio de los totales de los bloques en los que se aplica el tratamiento i. Asimismo, el estadístico $SS_{TRATAMIENTOS\ CORREGIDA}$ tiene $a-1$ grados de libertad. La descomposición de la variación total viene dada entonces por

$$SS_T = SS_{TRATAMIENTOS\ CORREGIDA} + SS_{BLOQUE} + SS_E.$$

Equivalentemente,

$$SS_E = SS_T - SS_{TRATAMIENTOS\ CORREGIDA} - SS_{BLOQUE}.$$

El estadístico entonces utilizado para contrastar la igualdad de los efectos de los tratamientos viene dado por

$$F_0 = \frac{MS_{TRATAMIENTOS\ CORREGIDA}}{MS_E},$$

donde las medias de cuadrados se calculan, como siempre, dividiendo por el número de grados de libertad ($a-1$ para $SS_{TRATAMIENTOS\ CORREGIDA}$ y $N-a-b+1$ para SS_E).

6.2. Estimación mínimo-cuadrática de los parámetros

Las ecuaciones normales de mínimos cuadrados para un modelo balanceado por bloques incompletos son:

$$\begin{aligned}
\mu:\ N\widehat{\mu} + r\sum_{i=1}^{a}\widehat{\tau}_i + k\sum_{j=1}^{b}\widehat{\beta}_j &= y_{..} \\
\tau_i:\ r\widehat{\mu} + r\widehat{\tau}_i + \sum_{j=1}^{b} n_{ij}\widehat{\beta}_j &= y_{i.}, \quad i = 1, \dots, a \\
\beta_j:\ k\widehat{\mu} + \sum_{i=1}^{a} n_{ij}\widehat{\tau}_i + k\widehat{\beta}_j &= y_{.j}, \quad j = 1, \dots, b.
\end{aligned} \tag{3}$$

Como soluciones de las anteriores ecuaciones, bajo las restricciones

$$\sum_i \widehat{\tau}_i = \sum_j \widehat{\beta}_j = 0,$$

se obtiene

$$\begin{aligned}
\widehat{\mu} &= \bar{y} \\
\widehat{\tau}_i &= \frac{kQ_i}{\lambda a}, \quad i = 1, \dots, a.
\end{aligned}$$

6.3. Recuperación de la información interbloques en un diseño balanceado

El análisis interbloques consiste en la obtención de información adicional sobre los efectos del tratamiento τ_i, cuando los efectos de bloque son variables aleatorias no correlacionadas con varianza σ_β^2.

El modelo para los totales por bloques, que configuran una muestra de b observaciones viene dado por

$$y_{\cdot j} = k\mu + \sum_{i=1}^{a} n_{ij}\tau_i + \left(k\beta_j + \sum_{i=1}^{a} \epsilon_{ij}\right),$$

donde el término entre paréntesis se interpreta como error. Los estimadores interbloques para μ y τ se obtienen minimizando la función

$$L = \sum_{j=1}^{b} \left(y_{\cdot j} - k\mu - \sum_{i=1}^{a} n_{ij}\tau_i\right)^2.$$

La solución de las ecuaciones normales derivadas de la función de pérdida anterior conducen a la siguiente solución:

$$\bar{\tau}_i = \frac{\sum_{j=1}^{b} n_{ij} y_{\cdot j} - kr\bar{y}_{\cdot\cdot}}{r-\lambda}, \quad i = 1, \ldots, a.$$

Tanto este estimador interbloques como el estimador obtenido anteriormente intrabloques $\widehat{\tau}_i$, $i = 1, \ldots, a$, son insesgados y sus varianzas vienen dadas por

$$\begin{aligned} V(\widehat{\tau}_i) &= \frac{k(a-1)}{\lambda a^2}\sigma^2 \\ V(\bar{\tau}_i) &= \frac{k(a-1)}{a(r-\lambda)}(\sigma^2 + k\sigma_\beta^2). \end{aligned}$$

Generalmente se consiera una combinación lineal de los estimadores intrabloques e interbloques

$$\tau_i^* = \alpha_1\widehat{\tau}_i + \alpha_2\bar{\tau}_i.$$

Los parámetros α_i, $i = 1, 2$, de la combinación lineal se determinan en términos de los parámetros σ^2 y σ_β^2 que serán estimados a partir de los datos. Específicamente, para $i = 1, \ldots, a$,

$$\tau_i^* = \frac{kQ_i(\sigma^2 + k\sigma_\beta^2) + \left(\sum_{j=1}^{b} n_{ij} y_{\cdot j} - kr\bar{y}_{\cdot\cdot}\right)\sigma^2}{(r-\lambda)\sigma^2 + \lambda a(\sigma^2 + k\sigma_\beta^2)}.$$