TEMA 3. Regresión lineal

1. Regresón lineal simple

1.1. Introducción

Los modelos de regresión surgen ante la necesidad de explorar la relación entre dos o más variables. En general, se suele identificar una o más variables de interés, que se denominan variables respuesta, cuyos valores son observados y se representan en términos de una función de una o más variables explicativas o independientes. La función que modeliza la relación entre la variable respuesta y las variables independientes o explicativas puede ser lineal o no lineal. Dado que usualmente las funciones continuas pueden ser aproximadas mediante funciones polinómicas, en este capítulo nos centraremos en el ajuste de modelos polinomiales. Es decir, para una variable respuesta Y se define la relación

$$Y = f(X_1, \dots, X_k),$$

donde, en este capítulo, f se supone una función polinómica, que relaciona los valores de la variables explicativas $X_1, \dots, X_k$ con los valores de la variable respuesta Y. Generalmente, los modelos de regresión son adecuados para el análisis de datos que provienen de experimentos que no fueron diseñados, o bien, en el caso de experimentos previamente diseñados, para establecer un modelo que relacione la variable respuesta o variables respuestas con los factores aleatorios previamente designados como importantes en el análisis de la varianza.

1.2. Regresión lineal simple

Se parte de una variable de regresión o explicativa X continua y controlable por el experimentador. En el diseño del experimento se determinan los valores de X que inducirán los correspondientes valores observados de la variable respuesta Y. Se supone una relación lineal entre al variable respuesta Y y la variable explicativa X. La observación de Y a cada nivel de X se modeliza mediate una variable aleatoria. Se tiene entonces que le valor esperado de Y para cada valor de X viene dado por

$$E\left[Y/X\right] = a_0 + a_1 X,$$

siendo a_0 y a_1 constantes desconocidas. Por tanto el modelo que define la observación de la variable respuesta Y viene dado por

$$Y = a_0 + a_1 X + \varepsilon,$$

representando ε la componente de error aleatoria. Específicamente, se supone que ε es una variable aleatoria con media cero y varianza σ^2 y que el conjunto de componentes aleatorias de error $\{\varepsilon\}$ no están correlacionadas. Partiendo de n pares de datos $(y_1, x_1), \ldots, (y_n, x_n)$, la estimación por mínimos cuadrados de los parámetros a_0 y a_1 del modelo se obtiene minimizando la función

$$L = \sum_{j=1}^{n} \varepsilon_i^2 = \sum_{j=1}^{n} (y_i - a_0 - a_1 x_i)^2.$$

Para minimizar dicha función se considera la siguiente expresión del modelo de regresión lineal:

$$Y = a_0' + a_1(X - \overline{X}) + \varepsilon,$$

siendo $a_0' = a_0 + a_1\overline{X}$, y $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$, que facilita la minimización de L y el cálculo de los estimadores $\hat{a}_0$ y $\hat{a}_1$ de los parámetros a_0 y a_1, cuyas expresiones, tras derivar L e igualar a cero, son

$$\begin{aligned} a_1 &= \frac{S_{XY}}{S_{XX}} \\ a_0 &= \overline{Y} - \hat{a}_1\overline{X}, \end{aligned} \tag{1}$$

siendo

$$\begin{aligned} S_{XY} &= \sum_{i=1}^{n} x_i y_i - n\overline{XY} \\ S_{XX} &= \sum_{i=1}^{n} x_i^2 - n\overline{X}^2 \end{aligned} \tag{2}$$

Los estimadores anteriormente calculados tienen las siguientes propiedades:

$$\begin{aligned} E[\hat{a}_1] &= a_1; \quad E[\hat{a}_0] = a_0 \\ \mathrm{Var}(\hat{a}_1) &= \frac{\sigma^2}{S_{XX}}; \quad \mathrm{Var}(\hat{a}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\overline{X}^2}{S_{XX}}\right] \end{aligned} \tag{3}$$

Para la estimación de la varianza σ^2 de la componente aleatoria de error, se utilizará la siguiente expresión de la suma de cuadrados de los residuos, SS_E :

$$SS_E = S_{YY} - \hat{a}_1 S_{XY},$$

de donde se obtiene, mediante el cálculo de la esperanza, que

$$E[SS_E] = (n-2)\sigma^2.$$

Por tanto, se considera como estimación de σ^2 la suma de cuadrados medios de los residuos definida como:

$$\hat{\sigma}^2 = \frac{SS_E}{n-2} = MS_E,$$

cuyo valor medio coincide con σ^2.

1.3. Inferencia en el modelo de regresión lineal simple

Contrastes de hipótesis asociados al modelo de regresión lineal simple

En la formulación de contrastes de hipótesis sobre el modelo de regresión lineal simple, se supondrá que los residuos se distribuyen normalmente. Se describen a continuación los estadísticos involucrados en el diseño de contrastes de ajuste sobre el modelo de regresión lineal simple.

Específicamente, en el ajuste de la pendiente de la recta, estaremos interesados en contrastar si el paparámetro a_1 toma un determinado valor a^*, es decir, se plantea el contraste

$$\begin{aligned} H_0: \ a_1 &= a^* \\ H_1: \ a_1 &\neq a^*. \end{aligned} \tag{4}$$

El estadístico formulado para el contraste de dicha hipótesis viene dado por

$$t_0 = \frac{\hat{a}_1 - a^*}{\sqrt{MS_E/S_{XX}}}.$$

Bajo la hipótesis nula, dicho estadístico sigue una distribución t-Student con $n-2$ grados de libertad. La hipótesis nula se rechaza cuando

$$|t_0| > t_{\alpha/2,n-2},$$

siendo $t_{\alpha/2,n-2}$ el percentil $1-\alpha/2$ de la distribución t-Student con $n-2$ grados de libertad.

Para el ajuste del orígen de la recta, dado por el parámetro a_0,

$$\begin{aligned} H_0: \ a_0 &= b^* \\ H_1: \ a_0 &\neq b^* \end{aligned} \tag{5}$$

se considera un estadístico similar definido por

$$t_0 = \frac{\hat{a}_0 - b^*}{\sqrt{MS_E\left(\frac{1}{n} + \frac{\overline{X}^2}{S_{XX}}\right)}}.$$

Nuevamente la hipótesis nula se rechaza si

$$|t_0| > t_{\alpha/2,n-2}.$$

Como caso especial, se puede considerar el contraste de ausencia de una relación lineal entre X e Y, o bien, la ausencia de una relación causal entre dichas variables, en términos del primer contraste de ajuste sobre la pendiente. Es decir,

$$\begin{aligned} H_0: \ a_1 &= 0 \\ H_1: \ a_1 &\neq 0 \end{aligned} \tag{6}$$

Mediante descomposición de la suma total de cuadrados corregida de Y,

$$S_{YY} = \sum_{i=1}^{n} (y_i - \overline{Y})^2 = SS_R + SS_E,$$

en términos de la variación de Y explicada por la recta de regresión y la variación residual, no explicada por la recta de regresión, se obtiene el diseño del estadístico

$$F_0 = \frac{SS_R/1}{SS_E/(n-2)} = \frac{MS_R}{MS_E},$$

para el contraste de la prueba de significación del modelo de regresión. Bajo la hipótesis nula, dicho estadístico sigue una distribución F-Snedecor con 1 y $n-2$ grados de libertad. Dicho estadístico puede ser derivado alternativamente considerando el cuadrado del estadístico utilizado en el contraste de ajuste del parámetro a_1, es decir, en términos del cuadrado de

$$t_0 = \frac{\hat{a}_1}{MS_E/S_{XX}},$$

que tras realizar unas pocas operaciones, se comprueba coincide con el cociente

$$\frac{MS_R}{MS_E},$$

que define el estadístico F_0.

Estimación por intervalos de confianza

En la estimación por intervalos de confianza de los parámetros que definen la recta de regresión se utilizan los estadísticos t-Student previamente formulados en los contrastes de ajuste descritos anteriormente. Más concretamente, a partir de los estadísticos

$$\frac{\hat{a}_1 - a}{\sqrt{MS_E/S_{XX}}}$$
$$\frac{\hat{a}_0 - b}{\sqrt{MS_E\left(\frac{1}{n} + \frac{\overline{X}^2}{S_{XX}}\right)}},$$

se obtienen los intervalos

$$\left[\hat{a}_1 \pm t_{\alpha/2,n-2}\sqrt{MS_E/S_{XX}}\right]$$
$$\left[\hat{a}_0 \pm t_{\alpha/2,n-2}\sqrt{MS_E\left(\frac{1}{n} + \frac{\overline{X}^2}{S_{XX}}\right)}\right]. \tag{7}$$

1.4. Idoneidad del modelo

Para comprobar la adecuación del ajuste por mínimos cuadrados, se puede realizar un análisis gráfico de los residuos, representando en papel normal una gráfica de los residuos frente a los valores ajustados y una gráfica de los residuos frente a cada variable de la regresión. También es útil la representación de dichos residuos frente a variables que no fueron incluidas en la regresión y que potencialmente pueden explicar parte de la variabilidad presentada por la respuesta.

Para comprobar la validez del modelo de regresión lineal simple también se utiliza, con frecuencia, un contraste sobre la idoneidad del modelo basado en el estadístico F

$$F_0 = \frac{MS_{LOF}}{MS_{PE}},$$

donde MS_{PE} denota la suma de cuadrados medios debida al error experimental y MS_{LOF} representa la suma de cuadrados medio debida a la falta de ajuste del modelo. Considerando m niveles de la variable explicativa X y n réplicas de la variable respuesta Y para cada nivel de la variable explicativa X se tiene, bajo la hipótesis nula de idoneidad del modelo, que el estadístico F_0 se distribuye según una $F-$Snedecor con $m-2$ y $n-m$ grados de libertad. Por tanto, la región crítica del test se define en términos de la cola superior de dicha F, es decir, la hipótesis de idoneidad del modelo se rechaza si

$$F_0 > F_{\alpha,m-2,n-m}.$$

Coeficiente de determinación

El coeficiente de determinación se define mediante la ecuación:

$$R^2 = \frac{SS_R}{S_{YY}} = \frac{\displaystyle\sum_{i=1}^{n}(\hat{y}_i - \overline{Y})^2}{\displaystyle\sum_{i=1}^{n}(y_i - \overline{Y})^2}.$$

Dicho coeficiente toma valores en el intervalo $(0, 1)$ y se interpreta como la proporción de variabilidad de los datos explicada por el modelo de regresión. Por este motivo, se suele utilizar R como un indicador de la adecuación del modelo de regresión, aunque se debe interpretar con cierto cuidado, pues, por ejemplo, R siempre aumenta de valor cuando se añade una nueva variable explicativa, lo que no supone que el modelo así obtenido sea superior al previo (con una variable menos). También hay que hacer notar que se puede conseguir que R tome el valor uno simplemente agregando un número suficiente de términos al modelo (tal es el caso del ajuste proporcionado por un polinomio de grado $n-1$ a una nube de n puntos).

2. Regresión lineal múltiple

2.1. Introducción

En el análisis estadístico y modelización de diversos procesos observables, representados mediante una variable respuesta, intervienen en su descripción más de una variable explicativa. Se adoptará la suposición de que el modelo que representa la variable respuesta en términos de las variables explicativas es lineal. Se estudiarán entonces los elementos que intervienen en la formulación de modelos de regresión lineal múltiple.

Se plantea, pues, el ajuste del modelo

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_k X_k + \epsilon.$$

Para la aproximación de Y a partir del hiperplano definido por la combinación lineal anterior, en el espacio $k-$ dimensional generado por las variables explicativas o variables de la regresión $X_1, \ldots, X_k$, se procederá a resolver el problema de estimación de los coeficientes de regresión $\beta_1, \ldots, \beta_k$.

2.2. Estimación puntual por mínimos cuadrados

Para la estimación mediante el método de mínimos-cuadrados del vector $(\beta_1, \ldots, \beta_k)$ se parte de n observaciones, suponiendo que $n > k$, así, mediante x_{ij} se notará la $j-$ésima observación o nivel de la variable X_i. La estimación puntual mínimo-cuadrática, se desarrollará, como en el modelo de regresión lineal simple, bajo la suposición de que las componentes aleatorias o componentes de error no estén correlacionadas, se hallen centradas y posean varianza σ^2. Considerando, para $j = 1, \ldots, n$, el modelo de observación

$$y_j = \beta_0 + \sum_{i=1}^{k} \beta_i x_{ij} + \epsilon_j,$$

se definirán las siguientes sumas de cuadrados

- *Suma de cuadrados corregida de la i-ésima variable de regresión*

$$S_{ii} = \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2 = \sum_{j=1}^{n} x_{ij}^2 - n\bar{x}_i^2, \quad i = 1, \ldots, k.$$

- *Suma corregida de los productos cruzados entre las variables X_r y X_s*

$$S_{rs} = \sum_{j=1}^{n} (x_{rj} - \bar{x}_r)(x_{sj} - \bar{x}_s).$$

- *Suma corregida de los productos cruzados entre X_i e Y*

$$S_{iy} = \sum_{j=1}^{n} y_j (x_{ij} - \bar{x}_i), \quad i = 1, \ldots, k.$$

En la derivación de los estimadores mínimo-cuadráticos de los coeficientes de regresión, se utilizará la transformación del modelo lineal de regresión múltiple dada por

$$\beta_0' = \beta_0 + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2 + \ldots \beta_k \bar{x}_k,$$

$$y_j = \beta_0' + \sum_{i=1}^{k} \beta_i (x_{ij} - \bar{x}_i) + \epsilon_j, \quad j = 1, \ldots, n.$$

Los estimadores de los parámetros se obtienen entonces minimizando la función cuadrática de pérdida dada por

$$L = \sum_{j=1}^{n} \left[y_i - \beta_0' - \sum_{i=1}^{k} \beta_i (x_{ij} - \bar{x}_i) \right]^2.$$

Equivalentemente, las ecuaciones normales que definen los estimadores mínimo-cuadráticos de los coeficientes de regresión se obtienen derivando parcialmente respecto a cada uno de los parámetros la función de pérdida L e igualando a cero:

$$n\widehat{\beta}_0' = \sum_{j=1}^{n} y_j$$

$$\sum_{l=1}^{k} \widehat{\beta}_l S_{il} = S_{iy}, \quad i = 1, \ldots, k.$$

La expresión matricial de las ecuaciones anteriores permite una resolución más directa y sencilla de las mismas, es decir, se parte de la ecuación matricial

$$\mathbf{y}_{n\times 1} = \mathbf{X}_{n\times k}\boldsymbol{\beta}_{k\times 1} + \boldsymbol{\epsilon}_{n\times 1},$$

siendo

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & (x_{11} - \bar{x}_1) & \ldots & (x_{k1} - \bar{x}_k) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (x_{1n} - \bar{x}_1) & \ldots & (x_{kn} - \bar{x}_k) \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0' \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

La función matricial de pérdida viene dada entonces mediante la expresión:

$$L = \mathbf{y}\mathbf{y}' - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}.$$

Derivando respecto al vector de parámetros $\boldsymbol{\beta}$ e igualando a cero se obtiene entonces

$$\mathbf{X}\mathbf{X}'\widehat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}.$$

Equivalentemente,

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}\mathbf{X}')^{-1}\mathbf{X}'\mathbf{y}.$$

A partir de la ecuación anterior se calcula la media y varianza del estimador mínimo-cuadrático del vector de parámetros que vienen dadas por

$$\begin{aligned} E[\widehat{\boldsymbol{\beta}}] &= \boldsymbol{\beta} \\ \mathrm{Cov}(\widehat{\boldsymbol{\beta}}) &= \sigma^2(\mathbf{X}\mathbf{X}')^{-1}. \end{aligned}$$

2.3. Contrastes de hipótesis en el modelo de regresión lineal múltiple

Para el planteamiento de contrastes de hipótesis sobre los parámetros del modelo de regresión múltiple se supondrá adicionalmente que la variable respuesta se distribuye según una normal. Más concretamente, las componentes aleatorias de error serán variables independientes e idénticamente distribuidas, según una normal con media cero y varianza σ^2. Por tanto, las variables y_j, para $j = 1, \ldots, n$, se distribuirán según una normal con media

$$\beta_0 + \sum_{i=1}^{k} \beta_i x_{ij}$$

y varianza σ^2.

Una de las hipótesis que usualmente se suelen contrastar en este contexto se refieren a la significación del modelo de regresión. Es decir, se suele contrastar

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = \cdots = \beta_k = 0 \\ H_1 &: \beta_i \neq 0, \quad \text{para algún } i. \end{aligned}$$

Rechazar la hipótesis nula significa que al menos una variable explicativa es necesaria o contribuye significativamente al ajuste. Similarmente, al caso de la regresión lineal simple, para deducir el estadístico F del test, se partirá de la descomposición de la suma total de cuadrados S_{yy} en términos de la suma de cuadrados de la regresión SS_R y la suma de cuadrados del error SS_E

$$S_{yy} = SS_R + SS_E.$$

Bajo la hipótesis nula,

$$SS_R/\sigma^2 \sim \chi^2(k),$$

es decir, la suma de suma de cuadrados de la regresión escalada mediante el inverso de la varianza poblacional se distribuye según una chi-cuadrado con k grados de libertad. También se tiene que

$$SS_E/\sigma^2 \sim \chi^2_{n-k-1}$$

y se tiene que SS_E y SS_R son independientes. Por tanto, se considera el estadístico

$$F_0 = \frac{SS_R/k}{SS_E/(n-k-1)} = \frac{MS_R}{MS_E},$$

que bajo la hipótesis nula se distribuye según una F de Snedecor, rechazándose la hipótesis nula si

$$F_0 > F_{\alpha,k,n-k-1}$$

siendo $F_{\alpha,k,n-k-1}$ el percentil $1-\alpha$ de la F de Snedecor con k y $n-k-1$ grados de libertad.

El cálculo de F_0 se realizará a partir de las siguientes identidades

$$\begin{aligned} SS_E &= S_{yy} - \sum_{i=1}^{k} \widehat{\beta}_i S_{iy} \\ SS_R &= \sum_{i=1}^{k} \widehat{\beta}_i S_{iy}. \end{aligned} \tag{8}$$

Otro aspecto interesante es contrastar si individualmente cada variable explicativa contribuye significativamente al ajuste del modelo de regresión. Dentro de éste ámbito se puede abordar también el problema de inclusión de nuevas variables explicativas que, en caso de ser necesarias, aumentarán la suma de cuadrados de la regresión y disminuirán las suma de cuadrados del error.

Para contrastar

$$\begin{aligned} H_0 &: \beta_i = 0 \\ H_1 &: \beta_i \neq 0 \end{aligned}$$

se considera el estadístico

$$t_0 = \frac{\widehat{\beta}_i}{\sqrt{MS_E C_{ii}}},$$

siendo C_{ii} el elemento i de la diagonal de la matriz $(\mathbf{X}'\mathbf{X})^{-1}$. El estadístico anterior se distribuye según una $t-$Student bajo la hipótesis nula, considerando que los estimadores de las componentes del vector de parámetros $\widehat{\beta}_i$, $i = 1, \ldots, k$, son independientes. Cuando la condición de independencia no se da, se pueden realizar conclusiones erróneas sobre la significación de los parámetros del modelo de regresión.

Para contrastar la significación de un conjunto de variables explicativas, se puede considerar el siguiente planteamiento: Se divide el vector de parámetros $\boldsymbol{\beta}$ en dos subvectores $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)^T$, siendo $\boldsymbol{\beta}_1$ un vector $r \times 1$ y $\boldsymbol{\beta}_2$ un vector $(p - r) \times 1$. Se desea contrastar entonces

$$\begin{aligned} H_0 &: \boldsymbol{\beta}_1 = 0 \\ H_1 &: \boldsymbol{\beta}_1 \neq 0 \end{aligned}$$

El modelo de regresión completo se puede expresar matricialmente como

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon},$$

denotando por $\mathbf{X}_1$ las columnas de $\mathbf{X}$ asociadas a $\boldsymbol{\beta}_1$ y $\mathbf{X}_2$ las columnas de $\mathbf{X}$ asociadas a $\boldsymbol{\beta}_2$.

Bajo la hipótesis nula, si se considera el modelo reducido

$$\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon},$$

se tiene que $\widehat{\boldsymbol{\beta}}_2 = (\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2\mathbf{y}$ y, por tanto,

$$SS_R(\boldsymbol{\beta}_2) = \widehat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{y},$$

con $p - r$ grados de libertad. Es decir, la suma de cuadrados de la regresión debida a $\boldsymbol{\beta}_1$ (o contribución de este conjunto de variables explicativas a la regresión), dado que ya se ha ajustado $\boldsymbol{\beta}_2$ (es decir, dada la contribución de las variables explicativas asociadas a $\boldsymbol{\beta}_2$) se define como

$$SS_R(\boldsymbol{\beta}_1/\boldsymbol{\beta}_2) = SS_R(\boldsymbol{\beta}) - SS_R(\boldsymbol{\beta}_2).$$

Bajo la hipótesis de que $SS_R(\boldsymbol{\beta}_1/\boldsymbol{\beta}_2)$ es independiente de MS_E, considerando la hipótesis nula se tiene que el estadístico

$$F_0 = \frac{SS_R(\boldsymbol{\beta}_1/\boldsymbol{\beta}_2)/r}{MS_E},$$

se distribuye según una $F-$Snedecor con r y $n - p$ grados de libertad. Se rechaza la hipótesis nula, es decir, se concluye que alguno de los parámetros de $\boldsymbol{\beta}_1$ es no nulo cuando $F_0 > F_{\alpha,r,n-p}$.

3. Modelo lineal general

Los resultados anteriores son válidos para cualquier modelo lineal en los parámetros $\beta_1, \ldots, \beta_k$, a ajustar. Por ejemplo, se puede considerar el efecto polinomial de un factor cuantitativo

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k.$$

También se puede considerar que las variables explicativas se definen en términos de un polinomio de segundo grado bivariante, es decir,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_2 x_2^2 + \beta_{12} x_1 x_2 + \boldsymbol{\epsilon}.$$

En ocasiones es interesante considerar un modelo trigonométrico

$$y = \beta_0 + \beta_1 \sin x + \beta_2 \cos x + \boldsymbol{\epsilon}.$$

En este contexto, también es factible el ajuste de modelos polinimiales ortogonales, es decir, de modelos definidos a partir de una combinación lineal de polinomios ortogonales aplicados a diferentes niveles de un factor que influye en la variable respuesta

$$y_j = \beta_0 + \beta_1 P_1(x_j) + \cdots + \beta_k P_k(x_j) + \boldsymbol{\epsilon}_j, \quad j = 1, \ldots, n,$$

siendo P_l un polinomio de grado l, satisfaciendo

$$\sum_{j=1}^{n} P_l(x_j) = 0, \quad \sum_{j=1}^{n} P_l(x_j) P_m(x_j) = 0, \quad l \neq m.$$

Bajo estas condiciones los resultados estadísticos anteriores se pueden derivar de forma similar en este contexto.