

Diseño de experimentos con medidas repetidas

Kevin García 1533173
Alejandro Vargas 1525953

7 de abril de 2019

Introducción

En esta presentación veremos el concepto y aplicación de los diseños de medidas repetidas, pasando por las ventajas y desventajas que este presenta y por los nuevos conceptos que surgen en este tipo de diseño, como los efectos de orden (efecto de período y efecto residual) y el supuesto de esfericidad, además se mostrarán los distintos tipos de diseño de medidas repetidas que podríamos tener (simple, factorial y mixto) y se realizará un ejemplo para ilustrar su aplicación.

Diseño de medidas repetidas

Cuando administramos los tratamientos objeto de investigación a los mismos sujetos y, en consecuencia, estos reciben más de un tratamiento experimental, disponiendo al menos de una observación por tratamiento y sujeto, decimos que estamos en presencia de un diseño intra-sujetos o de medidas repetidas. Es decir, todos los sujetos de la muestra reciben todos los tratamientos. De este modo, se trabaja con un solo grupo y cada individuo genera más de un dato. Estos diseños constituyen una alternativa válida y en muchos casos más adecuada que los diseños entre-sujetos o entre grupos.

Diseño de medidas repetidas

Estos diseños se pueden aplicar en diferentes tipos de situaciones, como:

- Evaluación longitudinal del cambio a lo largo del tiempo.
- Evaluación de la actuación de los sujetos bajo diferentes condiciones de tratamiento en estudios transversales.

Algunas de las ventajas del diseño de medidas repetidas son

- Reducción radical del error experimental. Esto es así porque la variabilidad debida a las diferencias individuales se elimina del término de error. De este modo, el diseño de medidas repetidas constituye una estructura más potente para detectar el efecto del tratamiento que los DCA (diseños completamente aleatorizados).

Diseño de medidas repetidas

- Menor costo en número de sujetos, la reducción del número de sujetos es una ventaja importante bajo múltiples punto de vista, incluidos el ahorro de tiempo y costos. Teniendo en cuenta que el mismo grupo de sujetos recibe todos los tratamientos.
- Incremento de la potencia del diseño. Es decir, dado que las diferencias entre los tratamientos son valoradas dentro de cada sujeto y no entre los grupos, la variabilidad atribuida a las diferencias individuales entre los sujetos, (efecto principal del factor sujetos), es eliminada del término de error, aumentando con ello precisión y potencia de la prueba estadística y acentuando así la validez de conclusión estadística (**Maxwell y Delaney, 1990; Vallejo, 1991**).

Diseño de medidas repetidas

Algunas de las desventajas de este tipo de diseño son

- El principal problema de los diseños de medidas repetidas son los efectos de orden que derivan de la propia estructura del diseño. Estos efectos deben ser neutralizados para que no confundan los efectos de los tratamientos. Pueden ser de dos tipos: efectos residuales y efectos de período.
- Los efectos de orden o la secuencia de administración de los tratamientos pueden sesgar los resultados. Una de las posibles soluciones ante este problema puede ser contrabalanceo equilibrando el orden de administración de los tratamientos mediante aleatorización de los tratamientos.

Diseño de medidas repetidas

- Otro inconveniente es la posible violación de algún supuesto de naturaleza estadística. Puesto que son los mismos sujetos los que reciben cada una de las condiciones experimentales es normal que en estos diseños aparezca un efecto sistemático en las respuestas, dando lugar a la aparición de correlación o dependencia entre los errores, algo por lo demás muy habitual en estos diseños (**Winer, 1971**). La violación de este supuesto repercute gravemente en los resultados obtenidos, ya que la prueba F no es robusta ante observaciones correlacionadas o dependientes.
- Otra desventaja que tienen estos diseños es que en la presencia de datos missing en alguna observación se deben tomar acciones: 1) Eliminar el sujeto del análisis: 2) Utilizar métodos de imputación.

Efectos de orden

Como se mencionó anteriormente, existen dos tipos de efectos de orden.

- Efecto de período: Los efectos de período ocurren cuando, independientemente del tratamiento aplicado, el sujeto responde al período o posición que, en la secuencia, ocupa el tratamiento (período de administración). Cabe, por lo tanto, la posibilidad de que el sujeto responda mejor al período que al tratamiento en sí mismo. Cuando esto ocurre, el efecto de período confunde la acción del tratamiento.

Solución: Contrabalanceo o aleatorización de los tratamientos.

Efectos de orden

- Efecto residual: El efecto residual, conocido por error progresivo, se caracteriza por la persistencia de la acción de un tratamiento más allá del período o tiempo de aplicación. Representa la progresiva acumulación tanto de los efectos facilitadores de la respuesta (efecto de la práctica, aprendizaje, etc.) como de los efectos obstaculizadores (como la fatiga mental, cansancio físico, etc.). Cuando, como es frecuente en esos casos, se produce una persistencia del efecto del tratamiento anterior sobre el tratamiento siguiente, se corre el riesgo de que los efectos queden contaminados.

Solución: Aumentar el intervalo de tiempo entre un tratamiento y el siguiente.

Supuestos

El adecuado análisis de los diseños de medidas repetidas requiere el cumplimiento de ciertos supuestos que aseguren la validez del contraste de hipótesis mediante la prueba F. Se asume la homogeneidad de las varianzas intra- tratamiento (esfericidad), y se asume que las observaciones en los diferentes niveles del tratamiento siguen una distribución normal conjunta multivariada. La violación de alguno de estos supuestos afectará al error de Tipo I y a la potencia de la prueba F; de entre ellos es especialmente grave la existencia de correlación entre los errores (**Winer, 1971; Kirk, 1982**).

Supuestos

El principal supuesto es el de esfericidad, que consta de homogeneidad o igualdad de las varianzas de las diferencias entre pares de puntuaciones. **(Huynt y Feldt 1970) y Rouanet y Lépine (1970)** demostraron que la homogeneidad de las varianzas de las diferencias equivale a asumir que la matriz de covarianza tiene determinada forma, denominada esfericidad o circularidad.

Existen procedimientos que permiten comprobar la hipótesis nula de esfericidad (la igualdad u homogeneidad de las varianzas poblacionales de las puntuaciones de diferencia). Una de ellas es la prueba W de Mauchly (1940) .

Clasificación

La clasificación del diseño de medidas repetidas en función de los grupos es:

- Diseño de medidas repetidas simple
- Diseño de medidas repetidas factorial
- Diseño de medidas repetidas factorial mixto

En esta presentación se analizará el diseño de medidas repetidas simple, ya que es el más sencillo de entender y a partir del análisis de este se podrían obtener los demás.

Diseño de medidas repetidas simple

La estructura del diseño de medidas repetidas simple es similar a la estructura de un diseño factorial de dos variables independientes. A diferencia del diseño factorial, la variable de sujetos no es manipulada ya que se trata de un pseudo-factor. La variable de tratamientos está manipulada por el experimentador y es considerada como un auténtico factor.

Diseño de medidas repetidas simple

Esta estructura está dada por:

$$y_{ij} = \mu + \tau_i + \alpha_j + \varepsilon_{ij} \quad i = 1, \dots, t \quad \wedge \quad j = 1, \dots, r$$

y_{ij} : Valor de la variable de respuesta del sujeto j bajo el tratamiento i .

μ : Media general sin tener en cuenta el sujeto ni el tratamiento.

τ_i : Efecto del i -ésimo tratamiento en la variable de respuesta.

α_j : Efectos del j -ésimo sujeto en la variable de respuesta.

ε_{ij} : Error aleatorio debido al i -ésimo tratamiento y al j -ésimo sujeto.

Supuestos:

$$\alpha_j \simeq N(0, \sigma_\alpha^2)$$

$$\varepsilon_{ij} \simeq N(0, \sigma_\varepsilon^2)$$

$$\Sigma = \sigma^2 11' + \sigma_\varepsilon^2 I$$

Análisis de varianza

En el análisis de varianza correspondiente para este tipo de diseño, se evalúan dos hipótesis esencialmente:

La hipótesis principal es la de los tratamientos

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_t = \tau$$

H_1 : Al menos una de estas igualdades no se cumple.

También existe interés en determinar si existen diferencias entre los sujetos

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_r = \alpha$$

H_1 : Al menos una de estas igualdades no se cumple.

Análisis de varianza

Las hipótesis anteriores se pueden probar realizando el análisis de varianza como sigue:

| Fuente de Variación | Grados de Libertad | Suma de Cuadrados | Cuadrados Medios | F | Pr > F | ECM |
|---------------------|--------------------|---|--|----------------------------------|--------|---|
| Tratamientos | $t - 1$ | $\sum_{i=1}^t \frac{y_{i.}^2}{r} - \frac{(y_{..})^2}{tr}$ | $CM_r = \frac{SC_r}{t-1}$ | $\frac{CM_r}{CM_{error}}$ | | $r \sum_{i=1}^t \frac{(\tau_i - \bar{\tau})^2}{t-1} + \sigma^2$ |
| Sujetos | $r - 1$ | $\sum_{j=1}^r \frac{y_{.j}^2}{t} - \frac{(y_{..})^2}{tr}$ | $CM_{\alpha} = \frac{SC_{\alpha}}{r-1}$ | $\frac{CM_{\alpha}}{CM_{error}}$ | | $t \sum_{j=1}^r \frac{(\alpha_j - \bar{\alpha})^2}{r-1} + \sigma^2$ |
| Error | $(t-1)(r-1)$ | $\vec{y}'\vec{y} - \sum_{i=1}^t \frac{y_{i.}^2}{r} - \sum_{j=1}^r \frac{y_{.j}^2}{t} + \frac{(y_{..})^2}{tr}$ | $CM_{error} = \frac{SC_{error}}{(t-1)(r-1)}$ | | | σ^2 |
| Total | $tr - 1$ | $\vec{y}'\vec{y} - \frac{(y_{..})^2}{tr}$ | | | | |

Cuadro: ANOVA

Ejemplo

Un total de 5 ratas son pesadas cuatro veces con intervalos de 4 semanas (semana 8, 12, 16 y 20). El objetivo es determinar el efecto del tiempo sobre el peso en gramos de las ratas. Los datos se presentan a continuación:

| Rata | Semana 8 | Semana 12 | Semana 16 | Semana 20 |
|------|----------|-----------|-----------|-----------|
| 1 | 164 | 220 | 261 | 306 |
| 2 | 164 | 230 | 275 | 326 |
| 3 | 158 | 226 | 264 | 320 |
| 4 | 159 | 227 | 280 | 330 |
| 5 | 155 | 222 | 272 | 312 |

Cuadro: Datos ejemplo

Ejemplo

- Factor: Tiempo
- Niveles: Semana 8, semana 12, semana 16 y semana 20
- Factor sujeto: Rata
- Variable de respuesta: Peso en gramos de las ratas

Modelo estadístico asociado:

$$y_{ij} = \mu + \tau_i + \alpha_j + \varepsilon_{ij} \quad i = 1, \dots, 4 \quad \wedge \quad j = 1, \dots, 5$$

y_{ij} : Peso en gramos de la rata j en la semana i .

μ : Peso promedio sin tener en cuenta las ratas ni las semanas.

τ_i : Efecto de la i -ésima semana en el peso de la rata.

α_j : Efecto de la j -ésima rata en el peso.

ε_{ij} : Error aleatorio debido a la rata j -ésima en la semana i -ésima.

Ejemplo

Supuestos:

$$\alpha_j \simeq N(0, \sigma_\alpha^2)$$

$$\varepsilon_{ij} \simeq N(0, \sigma_\varepsilon^2)$$

$$\Sigma = \sigma^2 11' + \sigma_\varepsilon^2 I$$

Descriptivas:

| Tiempo | \bar{X} | S | $CV(\%)$ | Min | Max |
|-----------|-----------|----------|----------|-------|-------|
| Semana 8 | 160 | 3.937004 | 2.46 | 155 | 164 |
| Semana 12 | 225 | 4 | 1.77 | 220 | 230 |
| Semana 16 | 270.4 | 7.829432 | 2.89 | 261 | 280 |
| Semana 20 | 318.8 | 9.859006 | 3.09 | 306 | 330 |

Cuadro: Descriptivas por tratamientos

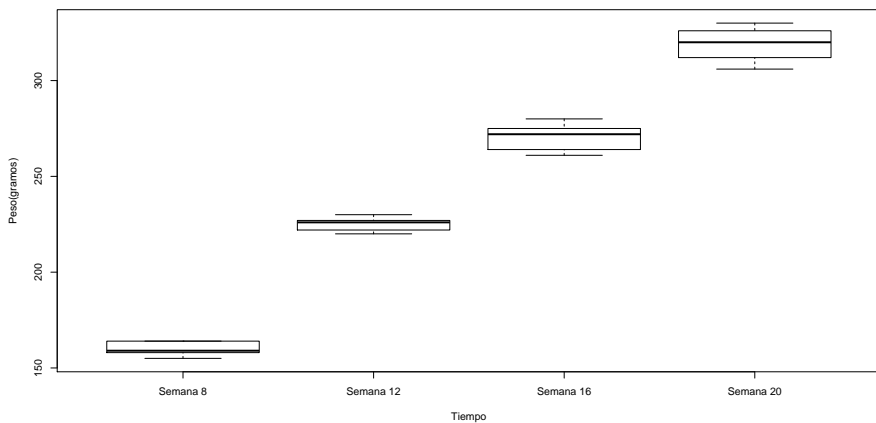
Ejemplo

| Tiempo | \bar{X} | S | $CV(\%)$ | Min | Max |
|---------------|-----------|----------|----------|-------|-------|
| Rata 1 | 237.75 | 60.42282 | 25.41 | 164 | 306 |
| Rata 2 | 248.75 | 68.77681 | 27.65 | 164 | 326 |
| Rata 3 | 242 | 68.01961 | 28.10 | 158 | 320 |
| Rata 4 | 249 | 73.27119 | 29.42 | 159 | 330 |
| Rata 5 | 240.25 | 67.71694 | 28.18 | 155 | 312 |

Cuadro: Descriptivas por sujetos

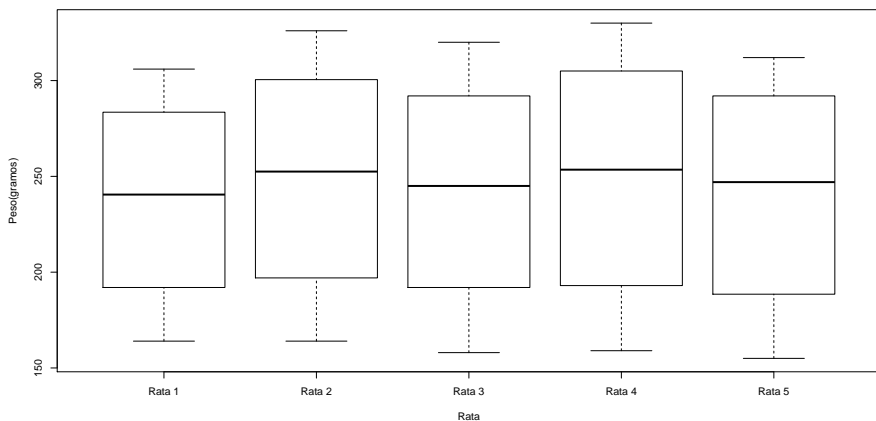
Ejemplo

Figura: Diagrama de cajas de los pesos(gramos) por semanas



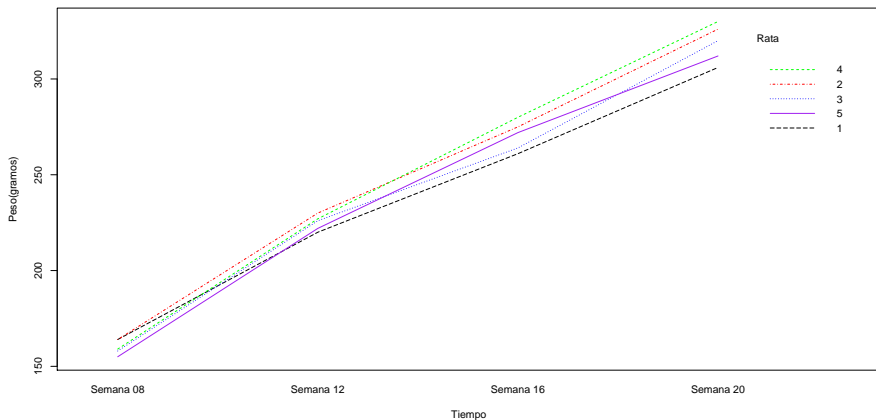
Ejemplo

Figura: Diagrama de cajas de los pesos(gramos) por rata



Ejemplo

Figura: Tendencia del peso por rata



Ejemplo

Primero verificamos los supuestos para poder realizar adecuadamente el análisis de varianza.

$$\alpha_j \simeq N(0, \sigma_\alpha^2)$$

Para validar este supuesto se utilizó la prueba Shapiro Wilk ya que se tienen muestras pequeñas.

Resultados:

Semana 8 **valor-p =0.359**

Semana 12 **valor-p =0.833**

Semana 16 **valor-p =0.7556**

Semana 20 **valor-p =0.8154**

A un nivel de significancia del 5 % No rechazamos la hipótesis nula para ninguno de los casos, es decir que hay evidencia muestral para pensar que se cumple la normalidad intra-sujetos.

Ejemplo

Ahora realizamos la prueba de esfericidad por medio del test de Mauchly para observar si hay o no esfericidad en la matriz Σ bajo las siguientes hipótesis.

H_0 : Hay esfericidad.

H_1 : No hay esfericidad.

Resultados:

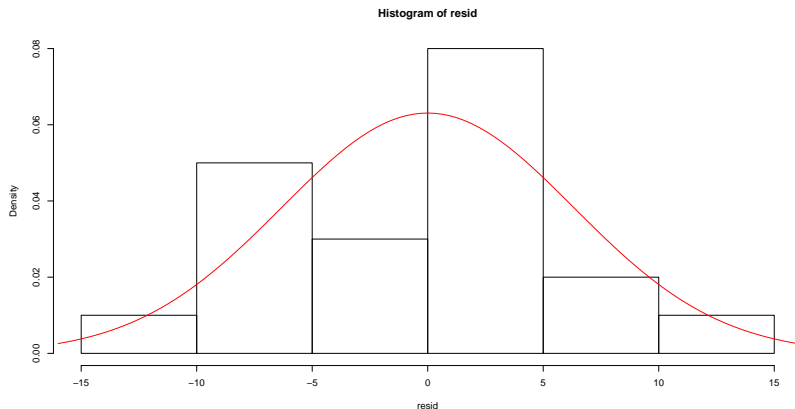
```
Mauchly Tests for Sphericity

          Test statistic p-value
design      0.15303 0.42942
```

A un nivel de significancia del 5 % no rechazamos H_0 , por lo tanto podemos concluir que se cumple el supuesto de esfericidad, es decir que no hay diferencias en las varianzas para todos los pares de grupos.

Ejemplo

Finalmente, se probó el supuesto de normalidad en los errores:



Y, la prueba Shapiro Wilk tuvo un $p\text{-valor} = 0.9801$.

Ejemplo

Luego de probar todos los supuestos, podemos confiar en que el análisis de varianza obtenido es válido.

Hipótesis

$$H_0 : \tau_1 = \tau_2 = \dots = \tau_t = \tau \quad \wedge \quad H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_r = \alpha$$

H_1 : Al menos una de estas igualdades no se cumple.

| Fuente de Variación | Grados de Libertad | Suma de Cuadrados | Cuadrados Medios | F | Pr > F | ECM |
|---------------------|--------------------|-------------------|------------------|-----|----------|---|
| Tiempo | 3 | 68541 | 22847 | 794 | 4.64e-14 | $r \sum_{i=1}^t \frac{(\tau_i - \bar{\tau})^2}{t-1} + \sigma^2$ |
| Rata | 4 | 414.7 | 103.7 | 3.6 | 0.03759 | $t \sum_{j=1}^r \frac{(\alpha_j - \bar{\alpha})^2}{r-1} + \sigma^2$ |
| Error | 12 | 345.3 | 28.775 | | | |
| Total | 19 | 69300.95 | | | | |

Cuadro: ANOVA Ejemplo

Podemos concluir que existen diferencias entre los efectos del tiempo (semanas) sobre el peso de las ratas a un nivel de significancia del 5 %.

Tópicos adicionales

Alternativas ante el incumplimiento de los supuestos

Cuando los supuestos mencionados no se satisfacen para las mediciones repetidas, los resultados del análisis descrito no son válidos, por lo que se sugieren varios análisis alternativos, ya que el análisis de varianza descrito anteriormente no se puede usar. Una de las alternativas hace referencia a ajustes conservadores del estadístico F usual del análisis de varianza para aproximar mejor los niveles de significancia de las pruebas, alternativa que se explica a continuación. (**Llobell, J. P., Navarro, M. D. F., & Pérez, J. F. G. (1996)**)

Tópicos adicionales

Para controlar este sesgo, bajo la hipótesis de nulidad de no diferencias entre las medidas repetidas, la distribución del estadístico F se puede aproximar a una distribución $F_{ajustado}$ con menos grados de libertad. El nivel de reducción o ajuste está dado por un parámetro ε que refleja la desviación de la matriz de varianzas y covarianzas poblacional de la homogeneidad.

| Grados de libertad F | F normal | F conservadora | F ajustada |
|----------------------|------------------|----------------|-----------------------------|
| Numerador | $t - 1$ | 1 | $\varepsilon(t - 1)$ |
| Denominador | $(r - 1)(t - 1)$ | $r - 1$ | $\varepsilon(t - 1)(r - 1)$ |

Cuadro: Ajuste de los grados de libertad

Una vez conocido el valor del factor correctivo (ε), simplemente hay que multiplicarlo por los grados de libertad originales del numerador y del denominador de la razón F y ajustar con ello los resultados para contrastar con las tablas de la distribución teórica de F .

Tópicos adicionales

El factor ε constituye un parámetro poblacional desconocido, que puede ser estimado a través de tres alternativas:

- Ajuste del límite inferior de Geisser-Greenhouse
- Ajuste de Box $\hat{\varepsilon}$
- Ajuste $\tilde{\varepsilon}$ de Huynh-Feldt

Estos métodos se encuentran implementados en los diferentes software estadísticos.

Referencias

- Kuehl, R. O. (2001), Diseño de experimentos
- Montgomery, D. C. (2004), Diseño y análisis de experimentos.
- Llobell, J. P., Navarro, M. D. F., & Pérez, J. F. G. (1996). Manual de psicología experimental.
- Arnau, J. (2007). Estudios longitudinales de medidas repetidas. Modelos de diseño y de análisis.
- Fox, J. & Weisberg, S. (2011). An R Companion to Applied Regression.

*<http://socserv.socsci.mcmaster.ca/~jfox/Books/Companion>