

Diseños con variables covariadas

Todos los diseños que se han presentado en la asignatura ya sean diseños *entresujetos* o *intrasujetos*, *unifactorial* o *multifactorial*, *aleatorizados* o *con restricciones*, pueden plantear hipótesis con una o más variables dependientes (*diseños multivariantes*, MANOVA, frente a los diseños *univariantes*, ANOVA, con una variable dependiente) y pueden también todos incluir alguna variable covariada cuantitativa continua que ajuste las puntuaciones de los sujetos y reduzca el término de error, mejorando la precisión en la estimación de los efectos experimentales, incrementando la potencia de la prueba estadística –aplicación básica dentro de los diseños propiamente experimentales–, o permita el control estadístico del efecto de ciertas variables perturbadoras conocidas como variables covariantes o covariadas, destacando como una técnica de análisis muy usual en los diseños cuasi-experimentales especialmente para controlar el sesgo de selección (Pascual, 1995c).

Por lo tanto, las variables covariadas representan a variables extrañas no directamente sometidas a investigación en la hipótesis pero cuyos efectos se desean controlar estadísticamente o su presencia puede deberse al objetivo de reducir el término de error. Su presencia en el estudio también constituye un identificador del rótulo del diseño de investigación experimental aplicado: ‘diseño de ANCOVA’.

- 1. Variables covariadas para reducir el término de error.** En los experimentos aleatorios (metodología experimental), donde los sujetos son aleatoriamente asignados a las condiciones experimentales, y si se cumplen los supuestos del ANCOVA, las diferencias esperadas entre las medias grupales respecto de la variable dependiente permanecerán siendo las mismas en un diseño de ANOVA y uno de ANCOVA, no disminuyendo la varianza entregrupos –efecto de los tratamientos– pero sí se reduce substancialmente el denominador de la prueba F que recoge la varianza de error. De este modo, la técnica del ANCOVA **incrementa la potencia del diseño experimental**.
- 2. Variables covariadas para controlar varianza sistemática secundaria (variables extrañas).** Cuando se plantea una variable como extraña en un determinado diseño es necesario controlarla con alguna técnica y una de ellas es el diseño con variables covariadas si la variable que se desea controlar se ha medido como una variable cuantitativa continua.

Cuando el diseño incluye alguna variable covariada se conoce como *diseño con covariables* o *diseño con variables covariadas* cuyo análisis se basa en la técnica del ajuste estadístico conocido como análisis de la covarianza, denominado ANCOVA cuando el diseño incluye una sola variable dependiente y MANCOVA cuando en el diseño se mide más de una variable dependiente. En el trabajo de Kirk (1995) por ejemplo se analizan cinco diseños con una o más variables covariadas: *diseño de análisis de la covarianza completamente aleatorizado*, *diseño factorial de análisis de la covarianza completamente aleatorizado*, *diseño de análisis de la covarianza de Cuadrado Latino*, *diseño de análisis de la covarianza de bloques completamente aleatorizados* y *diseño factorial de análisis de la covarianza split-plot*.

Por qué incluir una variable covariada en el diseño

El criterio básico que dirige la selección de la variable covariante es su alta correlación con la variable dependiente, de modo que cuanto mayor sea la correlación entre covariante y dependiente mayor error se sustraerá. Según Cox (1957) los beneficios derivados de la covariante **disminuyen** abiertamente cuando la correlación con la dependiente se sitúa por debajo de **.40**. **Cuando el diseño incluye más de una variable covariada ese mismo criterio se mantiene junto con la escasa o nula correlación entre las distintas covariantes ya que de lo contrario todas ellas sustraerán la misma varianza, no añadiendo ningún beneficio al diseño.** En la mayoría de los casos, el número de variables covariadas elegidas depende de criterios de utilidad, economía o tiempo, aunque algunos autores hacen depender la elección del tamaño de la muestra y del número de tratamientos administrados (Huitema,1980) mientras que otros autores incorporan los criterios de costo y beneficio (Allison, 1995).

El modelo estructural que describe al diseño entresujetos con una variable covariada es el siguiente:

$$\mathbf{Y} = \bar{\mathbf{Y}} + \mathbf{A} + b(\mathbf{X} - \bar{\mathbf{X}}) + \mathbf{E} \quad [1]$$

Donde:

- **Y** = puntuación en la variable dependiente
- **Y** = media de la muestra en la variable dependiente
- **A** = *efecto estimado* del tratamiento A en la variable dependiente

- b = pendiente de regresión Y sobre X
- \mathbf{X} = valores de la variable covariada
- $\bar{\mathbf{X}}$ = media de la variable covariada
- \mathbf{E} = error de estimación ajustado del modelo

En general se puede afirmar que el procedimiento del ANCOVA es robusto a las violaciones de la normalidad y de la homogeneidad de las varianzas tal y como ocurre en el ANOVA, sobre todo con igual número de observaciones por condición, pero ninguna de las dos técnicas de análisis es robusta a la violación del supuesto de independencia de las observaciones. Además, el ANCOVA exige el cumplimiento de otras condiciones, destacando la relación lineal entre la variable covariada y la variable dependiente, la independencia entre la variable covariada y la variable de tratamiento y la homogeneidad de la pendiente de regresión.

Es evidente que de **no haber relación entre las variables dependiente y covariada, la potencia del ANCOVA sería exactamente la misma que el ANOVA y, en consecuencia, la complejidad estadística añadida estaría carente de utilidad y sentido.**

También se exige la existencia de ortogonalidad entre las variables covariante e independiente, lo que equivale a suponer la no existencia de interacción entre las dos variables (Independiente y Covariada). La hipótesis de independencia es fácil de formular: equivale a asumir que *el tratamiento no afecta a la covariante*. Por tanto, si comprobamos mediante un simple ANOVA que la covariante presenta los mismos promedios en los distintos grupos de tratamientos habremos demostrado la independencia entre ambas.

Supongamos que se ha estimado el valor de la covariante para toda la muestra después que todos y cada uno de los sujetos han pasado por la situación de tratamiento experimental. Es plausible en estos casos asumir que el tratamiento experimental ha podido afectar a la variable dependiente y además a la variable covariante y, en consecuencia, al aplicar el ajuste estadístico de la una sobre la otra sustraeremos parte del efecto experimental generado por el tratamiento. El razonamiento es claro: si el tratamiento experimental afecta diferencialmente a la covariante, el ANCOVA, que iguala los grupos sobre la covariante, de hecho está “*removiendo*” varianza producida

por el tratamiento, sesgando por tanto la estimación de sus efectos. Como resultado de la reflexión anterior se impone la necesidad de medir la variable covariante siempre antes de administrar el tratamiento objeto de comprobación. Incluso, cumplido este requisito es aconsejable comprobar que la independencia se cumple.

Aunque la violación del supuesto de la homogeneidad de la pendiente de regresión no invalide la estimación estadística o la precisión de la misma, sí puede invalidar la interpretación de los datos obtenidos.

.....

Bloqueo y ANOVA

Si la variable de clasificación o control es medida cualitativamente, la técnica del análisis de bloqueo (*estratificación*) constituye una posible alternativa al procedimiento del ANCOVA. En opinión de Cochran (1957) las dos estrategias tienen prácticamente la misma eficacia en la reducción del error producido por una variable extraña si la correlación con dicha variable es de naturaleza lineal. Cox (1957) y Feldt (1958) han matizado un poco más su opinión, afirmando que el ANCOVA es superior al diseño de **bloques cuando la correlación covariante-dependiente es mayor que 0.80 y en cambio la estrategia de bloqueo es superior cuando la correlación es inferior a 0.40**. Por su parte, Maxwell y Delaney (1990) destacan la mayor sensibilidad del ANCOVA sobre el diseño de bloques, proponiendo que el diseño de bloques debe ser generalmente evitado en favor del diseño con variables covariadas.