LECCIÓN 6: CONTRASTES DE HIPÓTESIS.

SAAVEDRA, P.

1. Certezas y grados de evidencia

Comenzamos esta lección preguntándonos si en las biociencias pueden establecerse certezas absolutas. Cuando hacemos una revisión sistemática de una cuestión específica, a menudo podemos observar que los diversos estudios encontrados llegan a conclusiones diferentes y a veces, contradictorias. ¿A qué puede atribuirse esta heterogeneidad de resultados? Posiblemente a la ***complejidad*** de los objetos con los que operan las biociencias. Sin duda, los seres vivos obedecen a leyes universales para toda la naturaleza, pero el ser vivo es un *sistema abierto*, en el que intervienen múltiples variables de las que solo podemos observar unas pocas. Así por ejemplo, si queremos hacer una predicción del tiempo de supervivencia de un paciente a quien se le ha diagnosticado un enfermedad de muy mal pronóstico, sólo podremos hacerlo en términos de probabilidades debido a que las variables que podemos medir u observar no transportan más que una parte de la información sobre el estado real del sujeto.

Normalmente, la investigación científica parte de hipótesis basadas en el conocimiento apriorístico del sistema analizado, hipótesis que precisan ser probadas utilizando datos que contengan información sobre el referido sistema. Si los datos se contradicen con la hipótesis contraria a la considerada, los datos prevalecen y el consiguiente rechazo de la hipótesis contraria lleva al establecimiento de la hipótesis original como un nuevo hallazgo científico. ¿Pero, tal hallazgo podría considerarse como una certeza absoluta? A lo largo de esta lección veremos que en general no es posible establecer una certeza absoluta, sólo podremos hablar *grados de evidencia*. Orientamos esta lección a la medida del grado de evidencia a través del concepto de p-valor. Para entender el significado de este concepto debemos introducir previamente la idea de *contraste de hipótesis*. Presentaremos esta idea con el ejemplo que se muestra en la sección siguiente.

2. Asociación del polimorfismo $-842G > C$ en el gen PIN1 con el riesgo de cáncer de mama en mujeres de menos de 55 años

La enzima *PIN1* codificada por el gen *PIN1* es una isomerasa que causa cambios conformacionales en las proteínas e interviene en el crecimiento de las células epiteliales mamarias tanto fisiológica como patológicamente. Por tanto, las variantes genéticas del gen PIN1 pueden alterar la función de las proteínas y el riesgo de cáncer mamario. En tal sentido, Han *et al* (2010) desarrollaron un estudio de caso-control que puede resumirse del siguiente modo:

- *Objetivo del estudio.* Demostrar que el polimorfismo $-842G > C$ en el gen *PIN1* se asocia con una reducción del riesgo de cáncer de mama.

- *Diseño.* Estudio de *caso-control* en el que se incluyeron 467 mujeres con cáncer primario de mama y 488 controles, todas con 55 o menos años. Para este tipo de estudio, entenderemos que la asociación existe cuando la probabilidad de que una mujer tenga la mutación depende de que tenga o no el cáncer.

- *Resultados principales.* En la tabla 1 se resumen los principales resultados del estudio. Nótese que las edades en los grupos de casos y controles fueron similares. Para el polimorfismo $-842G > C$, la mutación apareció en el 23.3% de mujeres con cáncer y en el 31.1% en los controles. Estas cantidades *son sólo estimaciones* de la probabilidades de mutación en el grupo de mujeres con cáncer y en el grupo control respectivamente. A lo largo de esta sección analizaremos la idea de que existe *asociación estadísticamente significativa* entre la referida *mutación* y el cáncer de mama. El *grado de evidencia* viene medido por el llamado p-valor (0.007). ¿Qué es realmente esta cantidad?

En relación con el polimorfismo $-842G > C$, los datos sugieren la siguiente pregunta: ¿difieren las probabilidades de mutación $-842G > C$ entre las pacientes con cáncer y las de control? La *hipótesis* de asociación entre la *referida mutación* y el *riesgo de cáncer de mama*, sugerida por el conocimiento disponible acerca de la acción de la proteína codificada por el gen PIN1, podrá considerarse cierta cuando se demuestre usando los datos presentados. ¿*Cómo es el proceso de demostración*?

Pasamos ahora a exponer algunas ideas acerca de este proceso.

1. Representaremos por π_E y π_C las probabilidades de la mutación -842G>C en los casos (E) y controles (C) respectivamente.

TABLA 1. Estudio de Han et al (2010). Los datos son frecuencias absolutas y porcentajes por grupo de estudio

Variable	Casos ($n = 467$)		Controles ($n = 488$)		p-valor
	n	%	n	%	
Edad (años)					0.796
≤ 45	199	42.6	212	43.4	
$46 - 55$	268	57.4	276	56.6	
PIN1 -842G>C					**0.007**
GG	358	76.7	336	68.9	
*GC*ó *CC*	109	23.3	152	31.1	
PIN1 -667T>C					0.968
TT	189	40.5	194	39.8	
TC ó *CC*	278	59.5	294	60.2	

2. La hipótesis cuya verosimilitud se quiere evaluar es: $H_1 : \pi_E \neq \pi_C$. Esta hipótesis *que estaría describiendo una novedad científica* recibe el nombre de *hipótesis alternativa*.

3. Cabría pensar que la demostración se basaría en su compatibilidad con los datos disponibles para este fin.

4. Pero, ¿y si los datos también son compatibles con la hipótesis contraria H_0 (*hipótesis nula*)? ¿No sería mejor probar que los datos son ***incompatibles*** con H_0? Cabe decir que ***los datos prevalecen*** sobre las hipótesis que se formulen.

3. ELEMENTOS DE UN PROBLEMA DE CONTRASTE DE HIPÓTESIS

En general, un ***contraste de hipótesis*** es un *problema de decisión* entre una hipótesis nula H_0 y otra alternativa H_1 en un ambiente de incertidumbre; esto es, ***desconociéndose*** cuál es la cierta. Para realizar el contraste, el decisor debe obtener datos que le orienten a tomar la ***mejor decisión*** (aceptar o rechazar H_0). Basándonos en las las ideas examinadas en la sección anterior, describimos ahora los *elementos de un problema de contraste de hipótesis*, ilustrándolo con el estudio de asociación entre el polimorfismo -842G>C y el cáncer de mama (Han *et al* (2010)) descrito en la sección 2.

1. *Formulación de las hipótesis nula y alternativa.* (En el estudio de Han *et al*, $H_0 : \pi_E = \pi_C$ frente a $H_1 : \pi_E \neq \pi_C$).

2. *Datos con información acerca de las hipótesis descritas.* (En el estudio de Han *et al*, los que aparecen en la tabla 1 para el polimorfismo -842G>C). En general, son el resultado de la observación de *variables aleatorias*.

3. *Medida de discrepancia* T entre los datos y la hipótesis nula. Esta medida recibe el nombre de ***test estadístico***. (Para el, estudio de Han propondremos el test estadístico χ^2). Nótese que por depender de los datos, T es una ***variable aleatoria***.

4. *Regla de decisión.* Rechazar la hipótesis nula H_0 cuando el test estadístico T supere un cierto umbral C (a menudo, llamado valor crítico). Más concretamente y según sea el problema, cuando $T > C$ ó $\mid T \mid > C$.

Surge entonces la cuestión acerca de cuál debe ser el umbral crítico C a partir del que se ha de rechazar H_0. Para responder a esta pregunta, debemos considerar que la naturaleza elige si H_0 es cierta o falsa y el investigador o decisor, si la acepta o la rechaza. Los escenarios que pueden darse se resumen en la tabla 2.

TABLA 2. El contraste de hipótesis es un juego entre la naturaleza y el científico que debe decidir en ambiente de incertidumbre

		Decisión sobre H_0	
		Aceptarla	*Rechazarla*
H_0	*Cierta*	Decisión correcta	Error de tipo *alpha*
	Falsa	Error de tipo *beta*	Decisión correcta

Un contraste de hipótesis se dice que es ***estadísticamente significativo*** cuando la decisión que procede es **rechazar** H_0, y de esta forma, implementar una novedad científica.

4. El error de tipo *alpha*

En el escenario de un contraste de hipótesis, el error de tipo *alpha* se produce cuando se *rechaza una hipótesis nula cierta.* El modo natural de medir el error *alpha* es:

$$\alpha = \Pr\left(Rechazar\, H_0 \mid H_0\, es\, cierta\right)$$

En la investigación científica, el error *alpha* debe evitarse o bien, minimizarse. Normalmente, la hipótesis alternativa H_1 describe una novedad científica, pero si H_0 fuese cierta tal novedad sería una falacia. De esta forma, el rechazo de H_0 equivaldría a la *aceptación de una falsedad científica.*

Supóngase por ejemplo que en un ensayo clínico se establece como *hipótesis alternativa* H_1 que una cierta intervención terapéutica es superior a un placebo, frente a la *hipótesis nula* H_0 que indica que tal intervención es tan eficaz como el placebo. Cometer un error *alpha* significaría a establecer como eficaz la intervención terapéutica cuando realmente no lo es. Ello conduciría a una situación indeseable en la que se implementarían tratamientos ineficaces, que además podrían tener efectos adversos y el consiguiente coste económico.

En la sección anterior se planteó el problema de cuál debe ser el umbral para rechazar H_0. En la práctica, este problema se resuelve fijando a priori cual es el error *alpha* (α) que se está dispuesto a aceptar. Entonces, el valor crítico C se obtiene como solución de la siguiente ecuación:

$$\alpha = \Pr\left(T > C \mid H_0\, es\, cierta\right)$$

suponiendo que la regla de decisión consiste en: *rechazar* H_0 cuando $T > C$. Obviamente, cuando la regla sea: *rechazar* H_0 cuando $\mid T \mid > C$, el valor de C se establece como la solución de la ecuación:

$$\alpha = \Pr\left(\mid T \mid > C \mid H_0\, es\, cierta\right)$$

Nótese que para poder resolver las ecuaciones anteriores se precisa conocer la distribución del test estadístico T en el supuesto de que la hipótesis nula H_0 sea verdadera. Obsérvese que $\Pr\left(T \leq C \mid H_0\, es\, cierta\right)$ es precisamente su función de distribución para el valor C. Estas ecuaciones serán resueltas en las siguientes secciones para el test de la χ^2 y los test de la *t de Student* ó *t-test.*

5. Test χ^2 de Pearson para tablas 2×2

En esta sección se introduce el test de la χ^2 de Pearson para tablas de contingencia 2×2 ilustrado mediante el problema de la asociación del polimorfismo -842G>C con el cáncer de mama. En la tabla 3 se muestra una tabla de contingencia 2×2 correspondiente a la presencia/ausencia de la mutación en los grupos de casos (E) y controles (C).

Tabla 3. Estudio de Han et al (2010): los datos son frecuencias absolutas (%)

-842G>C	Cáncer (E)	Control (C)	Total
GC ó CC	**109 (23.3)**	**152 (31.1)**	261
GG	**358 (76.7)**	**336 (68.9)**	694
Total	467	488	955

Se plantea entonces el problema de ***decidir*** si existe asociación entre el polimorfismo -842G>C y el riesgo de cáncer de mama utilizando los datos de esta tabla. Tal asociación existe cuando $\pi_E \neq \pi_C$, siendo π_E y π_C las probabilidades de la mutación -842G>C en los grupos E y C respectivamennte. La hipótesis de *novedad científica*, y por tanto la *hipótesis alternativa* H_1 sería: $\pi_E \neq \pi_C$. Como se indicó en la sección 2, H_1 se establecerá como válida cuando los datos (resumidos en la tabla 3) ***se opongan*** a la hipótesis contraria (*nula*) $H_0 : \pi_E = \pi_C$. Se precisa por tanto definir una *medida de discrepancia* T (test estadístico) entre los datos y H_0. Para tal fin, proponemos el test de la χ^2 de Pearson para tablas 2×2.

En general, la formulación del test de la χ^2 se basa en dos poblaciones E y C en las que un factor F puede estar presente ($F = 1$) ó ausente ($F = 0$) en cada uno de los sujetos del estudio. Representamos por π_E y π_C las probabilidades de que el factor esté presente en las poblaciones E y C respectivamente. En orden a contrastar la hipótesis $H_0 : \pi_E = \pi_C$ frente a $H_1 : \pi_E \neq \pi_C$ se seleccionan muestras aleatorias de las poblaciones E y C de tamaños n_E y n_C respectivamente. Para la población i (E ó C), representamos finalmente por $n_{i,1}$ el número de sujetos con el factor F y por $n_{i,0}$, el número de los que no tienen el factor. Estos datos se pueden resumir entonces tal como se muestra en la tabla 4.

TABLA 4. Los datos son frecuencias absolutas. Nótese que $n_1 = n_{E,1} + n_{C,1}$ y que $n_0 = n_{E,0} + n_{C,0}$

	E	C	Total
$F = 1$	$n_{E,1}$	$n_{C,1}$	n_1
$F = 0$	$n_{E,0}$	$n_{C,0}$	n_0
Total	n_E	n_C	N

En orden a construir la medida de discrepancia entre los datos y la hipótesis nula H_0, haremos las siguientes consideraciones:

1. Para el tamaño muestral n_E fijado en E, el número de sujetos con el factor F es una variable aleatoria $n_{E,1}$ con distribución de probabilidad *binomial* (n_E, π_E). Entonces, el número esperado de sujetos en E con el factor F es $E\left[n_{E,1}\right] = n_E\pi_E$ (esperanza de la distribución binomial).
2. En la expresión de $E\left[n_{E,1}\right]$ podemos sustituir el parámetro desconocido π_E por su estimación $n_{E,1}/n_E$ (método *plug-in*).
3. Sin embargo, si H_0 es cierta, la probabilidad de poseer el factor F es independiente del grupo (E ó C) y por tanto, una mejor estimación de π_E es n_1/N (proporción global de pacientes con el factor, independientemente del grupo).
4. De acuerdo con los puntos anteriores, el valor esperado de sujetos con el factor F en el grupo E bajo H_0 es: $e_{E,1} = n_E \times n_1/N$. Si se observa la tabla de contingencia, puede verse que, bajo H_0, el número esperado de sujetos con el factor en el grupo E es igual al total de la columna (n_E) multiplicado por el total de la fila (n_1), dividido por el total global (N). Esta regla es idéntica para las cuatro celdas de la tabla de contingencia.

TABLA 5. Frecuencias observadas y esperadas bajo H_0

	E	C	Total
$F = 1$	$n_{E,1}$ $(e_{E,1} = n_E \times n_1/N)$	$n_{C,1}$ $(e_{C,1} = n_C \times n_1/N)$	n_1
$F = 0$	$n_{E,0}$ $(e_{E,0} = n_E \times n_0/N)$	$n_{C,0}$ $(e_{C,0} = n_C \times n_0/N)$	n_0
Total	n_E	n_C	N

En la tabla 5 se muestran, para cada una de las celdas de la tabla 2×2, las frecuencias observadas ($n_{i,j}$) y las esperadas bajo H_0 ($e_{i,j}$). Las frecuencias observadas son los ***datos*** y las esperadas bajo H_0 representan a H_0. ¿Cómo medir entonces la discrepancia de los datos con H_0?

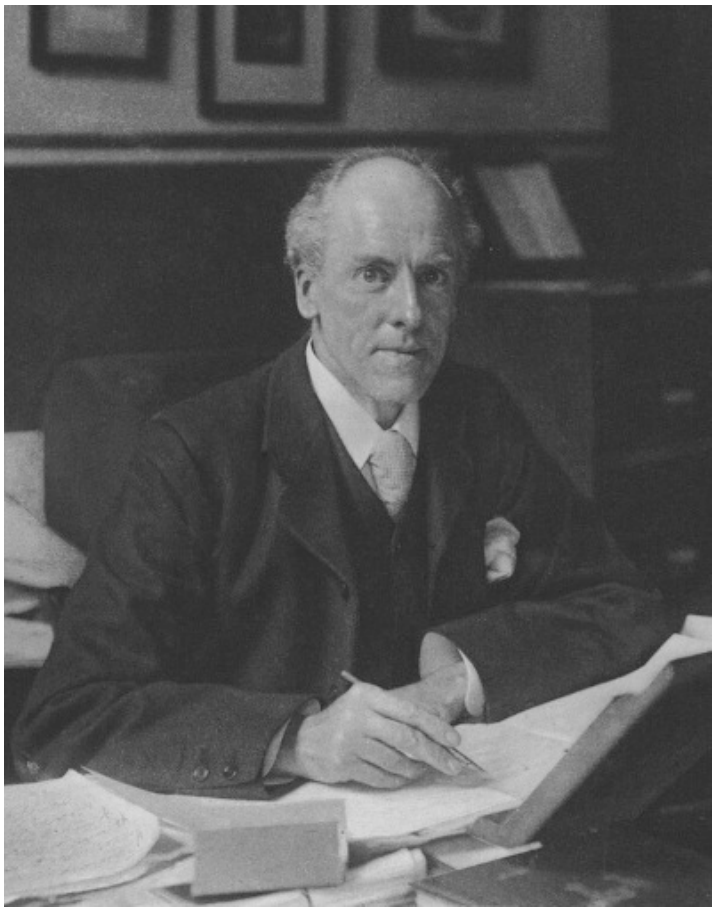

FIGURA 1. Karl Pearson (n. Londres 1857; m Londres 1936)

Karl Pearson (1857-1936) propuso en 1900 el siguiente *test estadístico* para medir la discrepancia entre los datos y H_0, a saber:

$$T = \sum_{i=E,C} \sum_{j=1,0} \frac{\left(n_{i,j} - e_{i,j}\right)^2}{e_{i,j}}$$

Una propiedad importante del test estadístico que se requiere para la determinación del valor crítico C la da la siguiente proposición.

Proposición: Bajo la hipótesis nula H_0, $T \cong \chi^2\,(1)$

Esto es, si la hipótesis nula H_0 es verdadera, el test estadístico T tiene distribución de probabilidad $\chi^2(1)$.

Finalmente, la regla de decisión del test χ^2 es: ***rechazar*** H_0, ***cuando*** $T > C$.

Para determinar el valor crítico C, fijamos el error *alpha* en la cantidad α y determinamos C como solución de la ecuación:

$$\alpha = \Pr\left(T > C \mid H_0 \; es \; cierta\right)$$

De acuerdo con la distribución de probabilidad de T bajo H_0 dada en la proposición anterior resulta obvio que: $C = \chi^2_{1-\alpha}(1)$ (cuantil 1-α de la distribución $\chi^2(1)$).

En la tabla tabla 6 se muestran las frecuencias observadas y esperadas bajo H_0 para el estudio de asociación entre el polimorfismo -2842G>C y el cáncer de mama.

TABLA 6. Estudio de Han *et al* (2010): los datos son frecuencias ***observadas*** y ***esperadas*** bajo H_0

-284G>C	Cáncer (E)	Control (C)	Total
GC ó CC	$n_{E,1} = 109$	$n_{C,1} = 152$	261
	$e_{E,1} = 127.6$	$e_{C,1} = 133.4$	
GG	$n_{E,0} = 358$	$n_{C,0} = 336$	694
	$e_{E,0} = 339.4$	$e_{C,0} = 354.6$	
Total	467	488	955

Se tiene entonces que el valor del test estadístico χ^2 es:

$$T = \frac{(109 - 127.6)^2}{127.6} + \frac{(152 - 133.4)^2}{133.4} + \frac{(358 - 339.4)^2}{339.4} + \frac{(336 - 354.6)^2}{354.6} \approx 7.323$$

Recuérdese que la regla de decisión correspondiente al test χ^2 es: ***rechazar*** H_0 ***cuando*** $T > \chi^2_{1-\alpha}$. ¿Qué procede hacer entonces? ¿Aceptar o rechazar H_0? La respuesta es clara, a saber: ***depende del error de tipo alpha dispuesto a asumir***.

En la tabla 7 se muestra la decisión que corresponde en función del error *alpha*.

TABLA 7. Nótese que dado que $T \approx 7,323$, procede rechazar H_0 cuando $7.323 > \chi^2_{1-\alpha}(1)$.

α	0.10	0.05	0.01	0.0068	0.005
$\chi^2_{1-\alpha}(1)$	2.71	3.84	6.635	7.3235	7.9794
Decisión	Rechazar H_0	Rechazar H_0	Rechazar H_0	Mínimo α de rechazo	No rechazar H_0

Nótese que el mínimo valor de α que permite rechazar H_0 es aproximadamente 0.007. Esta cantidad recibe el nombre de *nivel mínimo de significación* ó ***p*-valor**.

6. REDUCCIÓN MEDIANTE *EVOLOCUMAB* DE LOS NIVELES DE COLESTEROL EN PACIENTES CON SÍNDROME CORONARIO AGUDO

Para los pacientes con *síndrome coronario agudo* (SCA), las guías clínicas recomiendan el inicio hospitalario de terapia con *estatinas* con el objetivo de reducir los niveles de las *lipoproteínas de baja densidad* (C-LDL). Dado que este objetivo a menudo no se alcanza con las estatinas, se han buscado tratamientos alternativos a éstas. En tal sentido, Koskinas *et al* (2019) evaluaron la efectividad del ***evolocumab*** en combinación con la atorvastatina mediante un ensayo clínico controlado con placebo, aleatorizado y a doble ciego. Las características del ensayo se resumen en los siguientes puntos:

- *Objetivo del estudio.* Evaluar la viabilidad, seguridad y eficacia reductora del LDL-C de ***evolocumab*** iniciado durante la fase hospitalaria del SCA.
- *Métodos.* Ensayo clínico *controlado* con placebo, *aleatorizado* y a *doble ciego* en el que participaron 308 pacientes hospitalizados por SCA y con niveles elevados de C-LDL. Los pacientes incluidos fueron asignados al azar 1:1 para recibir 420 mg de evolocumab más atorvastatina (*tratamiento experimental*) o un placebo más atorvastatina (*tratamiento de control*). El criterio principal de valoración fue el *cambio porcentual en el C-LDL* entre el inicio y la octava semana.
- *Resultados.* Se resumen en la tabla 8. Cabe añadir que los eventos adversos y los eventos cardiovasculares fueron similares en ambos grupos.

En general, el cambio porcentual de una variable entre sus valores *basal* (X_{Basal}) y *final* (X_{Final}) se define por:

$$100 \times \frac{X_{Basal} - X_{Final}}{X_{Basal}}$$

TABLA 8. Los datos son medias ± SD y frecuencias (%). (*) Estos valores son independientes de las unidades en las que se mida el C-LDL. Se plantea como ejercicio obtener los p-valores que no aparecen en la tabla (señalados por ?)

	Evolocumab ($n = 155$)	Placebo ($n = 153$)	p-valor
Edad (años)	60.5 ± 12.0	61.0 ± 10.7	
Hombres	128 (82.6)	123 (80.4)	?
C-LDL (mmol/L)			
Basal	3.61 ± 1.00	3.42 ± 0.94	?
Semana octava	0.79 ± 0.46	2.06 ± 0.63	< 0.001
Reducción porcentual{*}	77.1 ± 15.8	35.4 ± 26.6	?

El objetivo pricipal del estudio es determinar si la reducción del C-LDL es superior en el brazo de tratamiento experimental (*evolocumab*) que en el de control (*placebo*). Denotamos entonces por μ_E y μ_C las reducciones porcetuales esperadas en los grupos experimental y de control respectivamente. Nuestra hipótesis alternativa es por tato: $H_1 : \mu_E > \mu_C$. Para este contraste utilizaremos en test de la t de student que se desarrolla en la siguiente sección.

7. Test de la t de Student (t-test)

Consideremos las poblaciones E (por ejemplo, todos los pacientes con SCA que reciben *evolocumab*) y C (por ejemplo, los pacientes con SCA que reciben el *placebo*). Representaremos por μ_E y μ_C los valores esperados (*esperanzas*) de un cierto marcador (por ejemplo, las reducciones porcentuales del C-LDL) en los brazos de tratamiento E y C respectivamente.

El primer elemento de un problema de contraste de hipótesis es su formulación. Según sea la naturaleza del problema, se considerarán los contrastes ***unilateral*** y ***bilateral,*** los cuales se definen en el modo que sigue:

1. *Contraste unilateral*
    1. Hipótesis nula. $H_0 : \mu_E = \mu_C$
    2. Hipótesis alternativa. $H_1 : \mu_E > \mu_C$
2. *Contraste bilateral*
    1. Hipótesis nula. $H_0 : \mu_E = \mu_C$
    2. Hipótesis alternativa. $H_1 : \mu_E \neq \mu_C$

Nótese que para la comparación de los valores basales del C-LDL en el estudio del evolocumab sería apropiado el contraste bilateral (los niveles basales del C-LDL deben ser similares) mientras que para comparar las reducciones porcentuales sería más apropiado el cotraste unilateral (el evolocumab es superior al placebo si reduce más el C-LDL).

Los contrastes anteriores se basarán en datos obtenidos de muestras aleatorias de las poblaciones E y C, de tamaños n_E y n_C respectivamente. Asumimos que los datos se extraen de distribuciones normales según el esquema que se muestra en la tabla 9.

TABLA 9. Distribución de la variable en cada población y datos que componen cada muestra.

Población	Distribución	Datos: muestras aleatorias
E	$N(\mu_E, \sigma_E)$	$X_{E,1}, ..., X_{E,n_E}$
C	$N(\mu_C, \sigma_C)$	$X_{C,1}, ..., X_{C,n_C}$

Inicialmente, la *medida de discrepancia* natural entre los datos y H_0 sería $\hat{\mu}_E - \hat{\mu}_C$, siendo, para $i = E, C$, $\hat{\mu}_i = (1/n_i) \sum_{j=1}^{n_i} X_{i,j}$ (estimadores para μ_E y μ_C respectivamente). En cualquier caso, $\hat{\mu}_E - \hat{\mu}_C$ depende de la unidades del marcador (en el ensayo del evolocumab, mmol/L). Dado que la medida de discrepancia debe ser independiente de las unidades de medida, consideramos finalmente el test estadístico (*adimensional*):

$$T = \frac{\hat{\mu}_E - \hat{\mu}_C}{\mathrm{sd}\,(\hat{\mu}_E - \hat{\mu}_C)}$$

De acuerdo con las propiedades de la varianza, es muy simple probar que: $\text{sd}\left(\hat{\mu}_E - \hat{\mu}_C\right) = \sqrt{\sigma_E^2/n_E + \sigma_C^2/n_C}$. Debido a que en general las varianzas σ_E^2 y σ_C^2 no se conocen, en la expresión de T son sustituidas por sus estimaciones $\hat{\sigma}_E^2$ y $\hat{\sigma}_C^2$ (método *plug-in*). De esta forma, la expresión final del test estadístico es:

$$T = \frac{\hat{\mu}_E - \hat{\mu}_C}{\sqrt{\hat{\sigma}_E^2/n_E + \hat{\sigma}_C^2/n_C}}$$

Según el test sea *unilateral* ó *bilateral*, consideramos las siguientes reglas de decisión:

1. *Contraste unilateral*: Rechazar H_0 cuando $T > C$.
2. *Contraste bilateral*: Rechazar H_0 cuando $\mid T \mid > C$.

De acuerdo con las ideas discutidas en las secciones anteriores, los valores críticos C se determinan a partir del error α que se asuma. Para ello se requiere conocer la distribución del test estadístico T bajo H_0, la cual se da de forma aproximada en la siguiente proposición.

Proposición: Bajo $H_0 : \mu_E = \mu_C$, $T \sim N(0,1)$

La distribución exacta de T es la llamada distribución t de Student, pero ésta es asintóticamente normal estándar ($N(0,1)$); esto es, que para grandes valores de n_E y n_C la distribución de T se comparta en la práctica como una $N(0,1)$.

De este resultado se siguen los corolarios 1 y 2, mediante los que se identifican los valores críticos para ambos contrastes.

Corolario 1: Para el contraste unilateral, $\Pr\left(T > z_{1-\alpha} \mid \mu_E = \mu_C\right) = \alpha$.

Este resultado indica que cuando se fija el *error alpha* en un valor α, el valor crítico del test ***unilateral*** es: $C = z_{1-\alpha}$ (ver figura 2)

Corolario 2: Para el contraste bilateral, $\Pr\left(\mid T \mid > z_{1-\alpha/2} \mid \mu_E = \mu_C\right) = \alpha$.

De forma análoga al corolario 1, este resultado indica que cuando se fija el *error alpha* en un valor α, el valor crítico del test ***bilateral*** es: $C = z_{1-\alpha/2}$ (ver figura 3)

Analizamos ahora la efectividad del *evolocumab+atorvastatina* (E) frente a *placebo+atorvastatina* (C) en la reducción de los niveles de C-LDL. Se trata por tanto de establecer si hay evidencias en que $\mu_E > \mu_C$ siendo μ_E y μ_C las reducciones esperadas

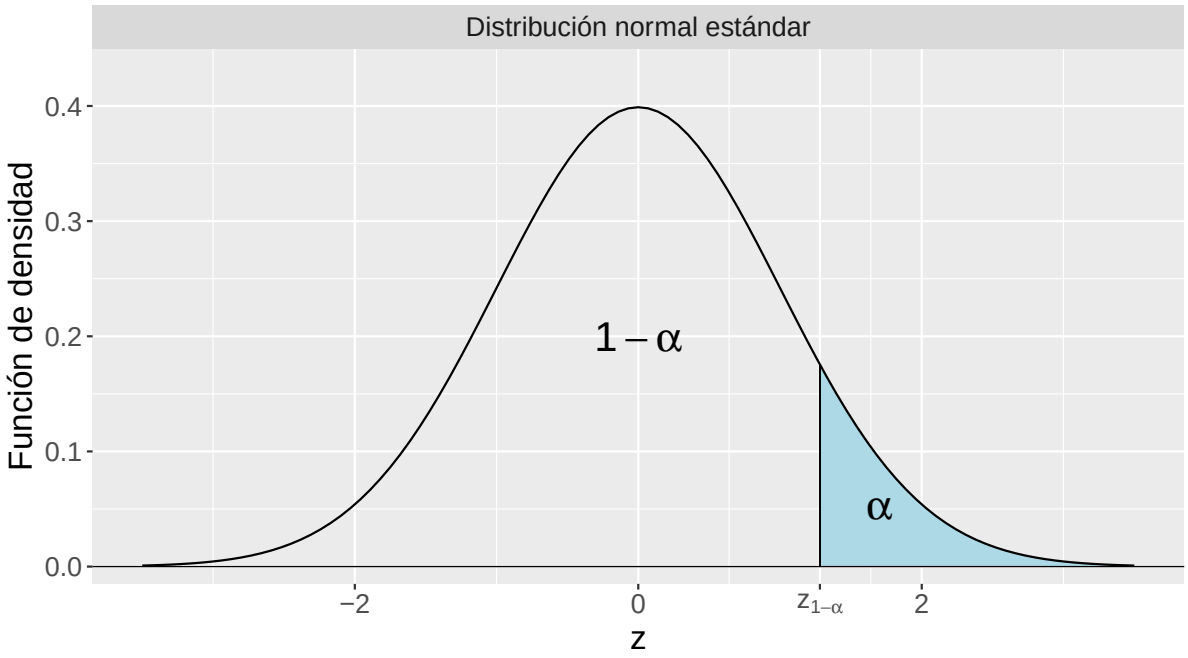

FIGURA 2. Distribución normal estándar. Para el test unilateral, el valor crítico de rechazo es $z_{1-\alpha}$ (cuantil $1-\alpha$ de la distribución normal estándar), el cual está determinado por el error α que está dispuesto a aceptar el decisor.

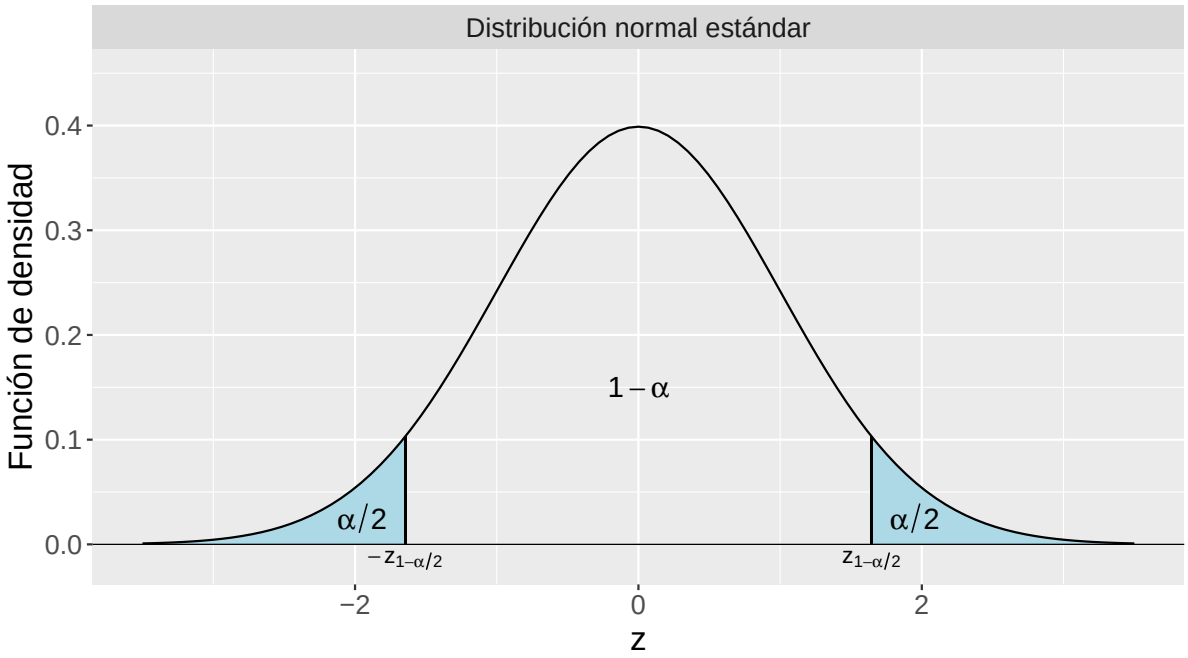

FIGURA 3. Para el test bilateral, la región de rechazo es la región marcada en azul

bajo los tratamientos E y C respectivamente. Si E no fuese efectivo, los valores esperados serían iguales. Por tanto, estaríamos ante un ***contraste unilateral*** de la forma: $H_0 : \mu_E = \mu_C$ frente a $H_1 : \mu_E > \mu_C$. Mediante los datos de la tabla 8 obtenemos el valor del test estadístico T como:

$$T = \frac{\hat{\mu}_E - \hat{\mu}_C}{\sqrt{\hat{\sigma}_E^2/n_E + \hat{\sigma}_C^2/n_C}} = \frac{77.1 - 35.4}{\sqrt{15.8^2/155 + 26.6^2/153}} \approx 16.7$$

Para este valor del test estadístico T, ¿se puede rechazar H_0? Como ya se ha señalado, *depede del error* α que se esté dispuesto a asumir. Más concretamente, para un valor del error α fijado, la regla de decisión es ahora: *Rechazar* H_0 cuando $T \approx 16.7 > z_{1-\alpha}$. En la tabla 10 se muestran para diferentes valores de α, los correspondientes valores críticos $z_{1-\alpha}$ y la decisión que procede tomar.

TABLA 10. Nótese que el p-valor es inferior a 0.00001; Hay una fortísima evidencia en contra de H_0

α	0.10	0.05	0.01	0.001	0.00001
$z_{1-\alpha}$	1.282	1.645	2.326	3.090	4.265
Decisión sobre H_0	*Rechazar*	*Rechazar*	*Rechazar*	*Rechazar*	*Rechazar*

El estudio del *evolocumab* es un *ensayo clínico aleatorizado*, lo que garantizaría que los grupos experimentales fuesen ***homogéneos*** antes de la aplicación de los tratamientos. Este es un aspecto esencial de los ensayos, pues permite que las eventuales diferencias entre los tratamientos puedan atribuise a sus efectos y no a factores de confusión. En este sentido cabría preguntarse: ¿Existen diferencias entre los valores esperados del C-LDL en la situación basal (antes de la aplicación de los tratamientos)? El ensayo no sería válido si tales diferencias fuesen significativas.

Para analizar el problema de la homogeneidad de los niveles basales del C-LDL, designamos por μ_E y μ_C los valores esperados del marcador en los grupos E y C respectivamente y consideramos el contraste bilateral: $H_0 : \mu_E = \mu_C$ frente a $H_1 : \mu_E \neq \mu_C$. El valor del test estadístico es ahora:

$$T = \frac{\hat{\mu}_E - \hat{\mu}_C}{\sqrt{\hat{\sigma}_E^2/n_E + \hat{\sigma}_C^2/n_C}} = \frac{3.61 - 3.42}{\sqrt{1^2/155 + 0.94^2/153}} \approx 1.718$$

Para un valor del error α fijado, la regla de decisión es ahora: *Rechazar* H_0 cuando $\mid T \mid \approx 1.718 > z_{1-\alpha/2}$. De la tabla 11 puede deducirse el p-valor del contraste.

TABLA 11. Nótese que el p-valor es 0.08574 dado que éste es el mínimo valor α que permite rechazar H_0

α	0.10	0.09	0.08574	0.085	0.05
$z_{1-\alpha 72}$	1.645	1.695	1.718	1.722	1.960
Decisión sobre H_0	*Rechazar*	*Rechazar*	*Rechazar*	*Aceptar*	*Aceptar*

De este contraste se deduce, que para admitir diferencias entre los niveles basales del C-LDL, habría que aceptar un error α superior a 0.08574, lo cual no es admisible en la innvestigación científica.

8. REFERENCIAS

1. Han CH, Lu J, Wei Q, Bondy ML, Brewster AM, Yu TK, Buchholz TA, Arun BK, Wang LE. The functional promoter polymorphism (-842G>C) in the PIN1 gene is associated with decreased risk of breast cancer in non-Hispanic white women 55 years and younger. Breast Cancer Res Treat. 2010 Jul;122(1):243-9. doi: 10.1007/s10549-009-0682-9. Epub 2009 Dec 24. PMID: 20033770; PMCID: PMC2883663. pdf

2. Koskinas KC, Windecker S, Pedrazzini G, Mueller C, Cook S, Matter CM, Muller O, Häner J, Gencer B, Crljenica C, Amini P, Deckarm O, Iglesias JF, Räber L, Heg D, Mach F. Evolocumab for Early Reduction of LDL Cholesterol Levels in Patients With Acute Coronary Syndromes (EVOPACS). J Am Coll Cardiol. 2019 Nov 19;74(20):2452-2462. doi: 10.1016/j.jacc.2019.08.010. Epub 2019 Aug 31. PMID: 31479722. pdf