

# **Análisis estadístico de las medidas de difusividad obtenidas a partir de muestras obtenidas mediante imagen de tensor de difusión**

Por: Daniela M. Martínez Plata

[daniela.martinez@uabc.edu.mx](mailto:daniela.martinez@uabc.edu.mx)

Facultad de Ciencias de la Ingeniería y Tecnología

Universidad Autónoma de Baja California

Valle de las Palmas, Baja California, México

## *Resumen*

El presente documento describe los resultados obtenidos al aplicar una metodología de análisis estadístico inferencial a un conjunto de datos muestrales obtenidos para cuatro medidas de difusividad extraídas de imágenes obtenidas mediante la técnica denominada tensor de difusión. Este análisis forma parte del proyecto denominado “Comportamiento de las medidas de difusividad en sujetos con riesgo genético para enfermedad de Alzheimer”.

## *Planteamiento del análisis*

A partir de la revisión del comportamiento teórico esperado para la anisotropía fraccional y la difusividad media, se tiene que la afectación de los tractos de la proteína TAU producen un incremento en la difusividad media y un decremento en la anisotropía fraccional de la sustancia blanca. Dado que la afectación a los tractos de la proteína se relaciona con la producción de neurotoxicidad por mecanismos de disfunción sináptica y estados excitotóxicos que inducen a la muerte neuronal, el estudio de este comportamiento podría utilizarse para el diagnóstico temprano de la enfermedad de Alzheimer.

El proyecto denominado “Comportamiento de las medidas de difusividad en sujetos con riesgo genético para enfermedad de Alzheimer” tiene como objetivo general analizar el comportamiento de las medidas de difusividad evaluadas por imagen de tensor de difusión en sujetos con riesgo genético para enfermedad de Alzheimer, para lo cual se recolectó una muestra de 218 sujetos, clasificados de la siguiente manera:

Tabla 1. Descripción de los grupos definidos para la muestra	
Grupo	Número de muestras
Demencia	49
Controles para demencia	30
Deterioro cognitivo leve (DCL)	38
Controles para DCL	42
Sanos portadores	35
Sanos no portadores	24

Como parte de los análisis realizados, se condujo un análisis de varianzas (ANOVA) para los distintos grupos considerados en la muestra, identificando que existe disminución en la anisotropía fraccional y aumento en la difusividad media, axial y radial en los portadores de la mutación E280A con respecto a los no portadores y que el mismo comportamiento se evidencia para los grupos de deterioro cognitivo leve y demencia con respecto a sus controles

En este sentido, se busca complementar los resultados del análisis de varianzas con una prueba paramétrica de hipótesis para validar que el promedio de la difusividad media del grupo de sujetos portadores de la mutación PSEN-1 E280A es significativamente mayor que el promedio de la difusividad media del grupo de sujetos no portadores de dicha mutación. Así mismo, se busca validar que el promedio de la anisotropía fraccional del grupo de sujetos portadores de la mutación PSEN-1 E280A es significativamente menor que el promedio de la anisotropía fraccional del grupo de sujetos no portadores de dicha mutación.

### *Metodología*

La primera etapa del análisis consiste en la validación de la normalidad de la distribución de las observaciones para cada una de las métricas; lo cual se validó a partir de la prueba de bondad de ajuste Kolmogorov-Smirnov (K-S). La prueba K-S corresponde a una prueba para la hipótesis de que los datos de una determinada muestra proceden de una distribución normal estándar, frente a la alternativa de que no procede de tal distribución [1].

En este caso, se procedió a estandarizar las métricas de anisotropía fraccional (AF), difusividad media (DM), difusividad axial (AD) y difusividad radial (RD) para cada uno de los grupos de muestras que se están analizando en el estudio, a partir del siguiente modelo:

$$Z_{K-S} = \frac{x_s - \bar{x}_s}{s_s} \quad (1)$$

En donde los valores de las observaciones para cada uno de los grupos, denotados por  $\mathbf{x}_s$ , se utilizan para construir un estadístico de prueba ( $z_{K-S}$ ), a partir de cual se efectúa una prueba de normalidad con la cual se puede afirmar con un  $(1 - \alpha)\%$  de confianza que el conjunto de datos muestrales ( $\mathbf{x}_s$ ) proviene de una distribución normal estándar, en donde  $\bar{x}_s$  y  $s_s$  corresponde al promedio y la desviación estándar del conjunto de observaciones de la métrica analizada, respectivamente.

Para el presente estudio se utilizaron las funciones proporcionadas por la *librería Statistics and Machine Learning* de Matlab, un software de cómputo numérico empleado para realizar las validaciones de los conjuntos de datos.

Los resultados arrojados por el software, como se muestran en las Figuras 1 a la 6, demuestran que todas las métricas de difusividad, para todos los grupos de estudio dieron positivo a la normalidad de la distribución. Es decir, la K-S demostró que los

datos para la AF, DM, AD y RD en todos los grupos de estudio presentan una distribución normal, considerando una confianza del 95%.

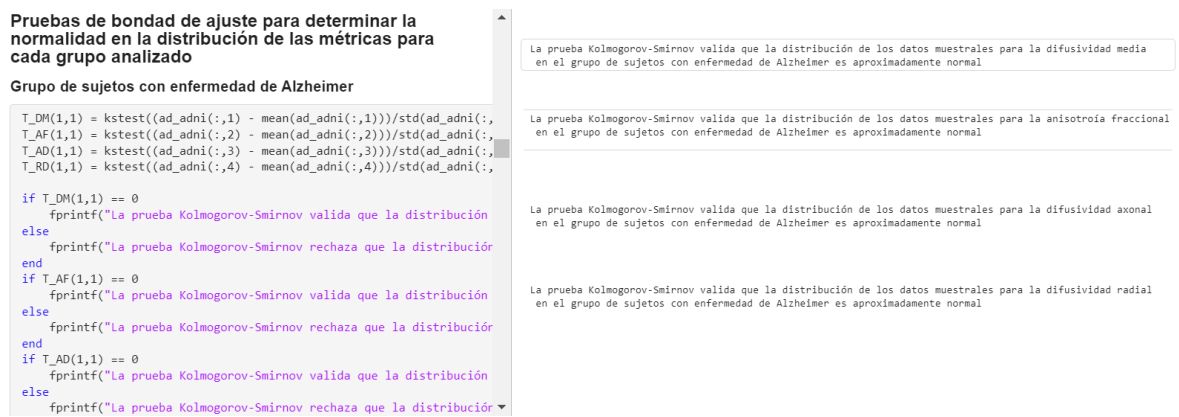


Figura 1. Resultados de la prueba K-S para el grupo de sujetos con enfermedad de Alzheimer.

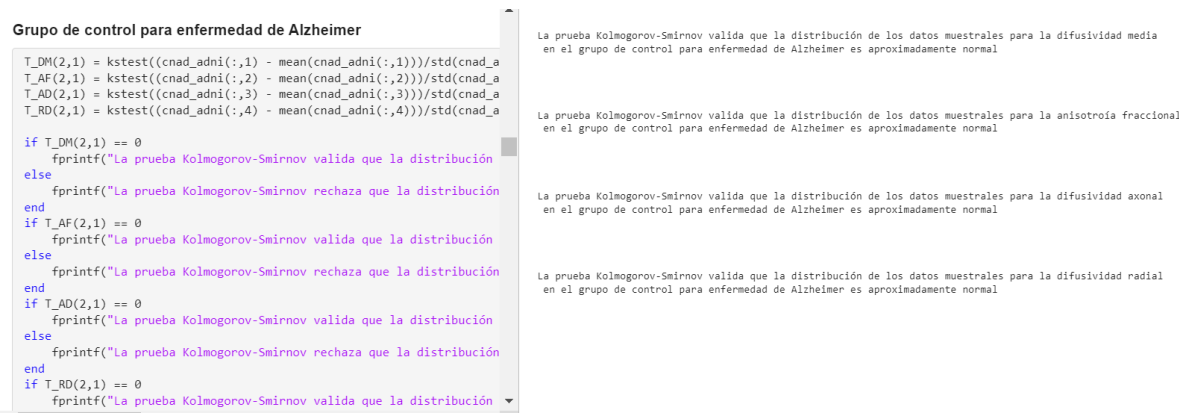


Figura 2. Resultados de la prueba K-S para el grupo de control para la enfermedad de Alzheimer.

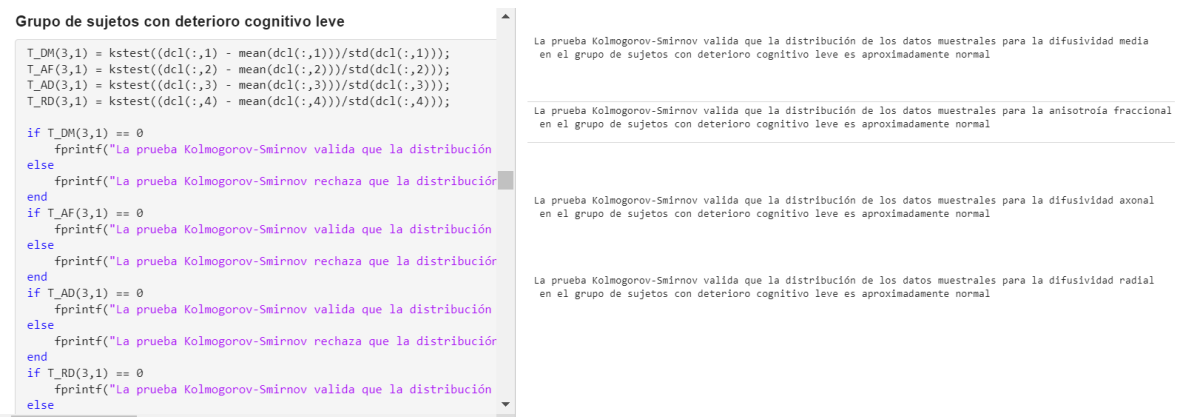


Figura 3. Resultados de la prueba K-S para el grupo de sujetos con deterioro cognitivo leve.

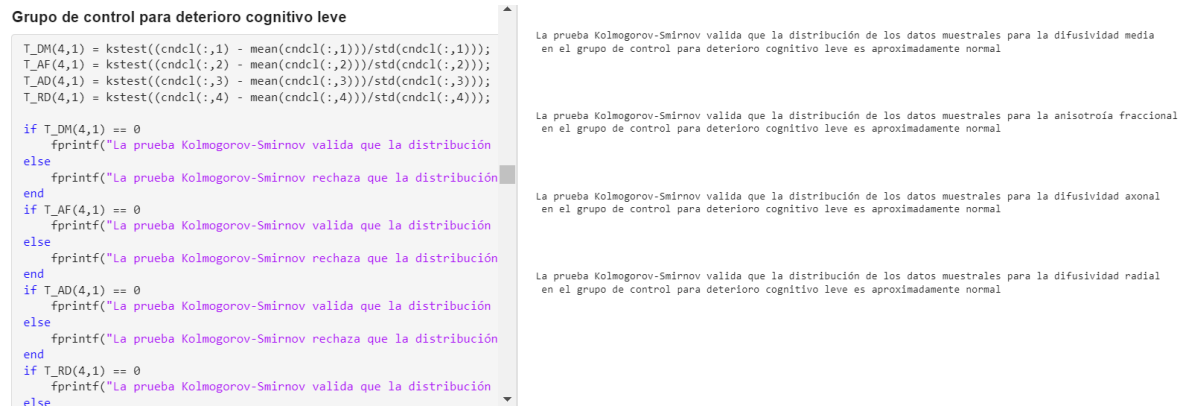


Figura 4. Resultados de la prueba K-S para el grupo de control para el deterioro cognitivo leve.

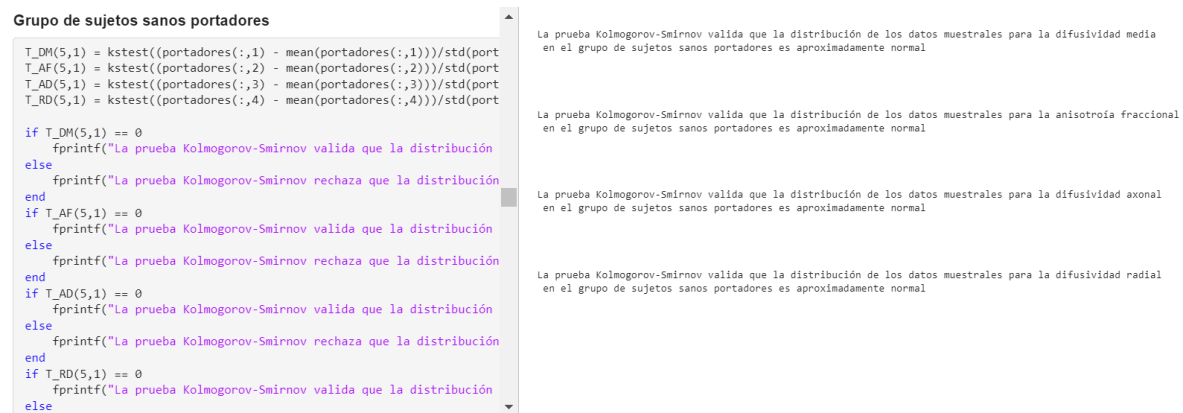


Figura 5. Resultados de la prueba K-S para el grupo de sujetos sanos portadores de la mutación del gen.

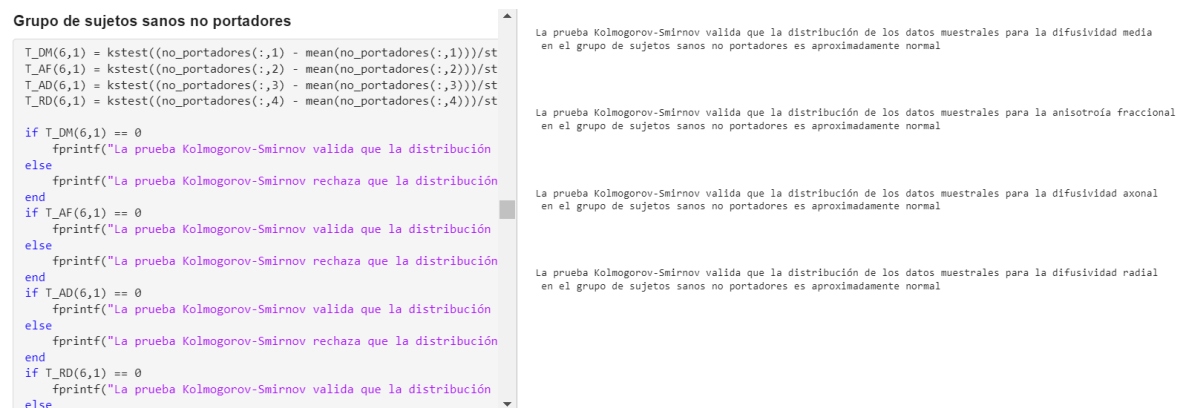


Figura 6. Resultados de la prueba K-S para el grupo de sujetos sanos no portadores de la mutación del gen.

Una vez que se demostró que la distribución subyacente para el conjunto de datos muestrales de cada grupo corresponde a una distribución aproximadamente normal, se procedió a la formulación de las pruebas de hipótesis paramétricas que validarán

que existen diferencias significativas entre la AF y la DM para el grupo de sujetos portadores y no portadores de la mutación E280A.

Para ello se formularon los dos siguientes conjuntos de hipótesis:

1.  $H_0: \mu_{AF,p} - \mu_{AF,np} = 0$   
 $H_1: \mu_{AF,p} - \mu_{AF,np} < 0$
2.  $H_0: \mu_{DM,p} - \mu_{DM,np} = 0$   
 $H_1: \mu_{DM,p} - \mu_{DM,np} > 0$

Para el conjunto de hipótesis (1) se establece la premisa que la diferencia entre los promedios de la anisotropía fraccional en sujetos portadores de la mutación E280A ( $\mu_{AF,p}$ ) y de los sujetos no portadores ( $\mu_{AF,np}$ ) debe ser menor que cero para comprobar la afectación de los tractos de la proteína TAU. Por otro lado, para el conjunto de hipótesis (2) se establece que el promedio de la difusividad media en sujetos portadores de la mutación E280A ( $\mu_{DM,p}$ ) debe ser menor que el promedio de la difusividad media en los sujetos no portadores de la mutación ( $\mu_{DM,np}$ ) para confirmar la afectación de los tractos de la proteína TAU.

La prueba que se debe realizar en el caso del conjunto de hipótesis 1 corresponde a un contraste unilateral de cola izquierda, como lo establece la siguiente expresión:

$$\begin{array}{c} H_0 \\ t_{\Delta\bar{x}_{AF}} > \lambda_{\alpha,AF} \\ H_1 \end{array} \quad (2)$$

Mientras que la prueba para el caso del conjunto de hipótesis 2 corresponde a un contraste unilateral de cola derecha, representado mediante:

$$\begin{array}{c} H_0 \\ t_{\Delta\bar{x}_{DM}} < \lambda_{\alpha,DM} \\ H_1 \end{array} \quad (3)$$

Los elementos que integran los contrastes (2) y (3) corresponden al estadístico de prueba calculado a partir del promedio de las observaciones de la muestra ( $t_{\bar{x}}$ ) y el valor crítico ( $\lambda_{\alpha}$ ) obtenido para el nivel de significancia ( $\alpha$ ) elegido para la prueba. Ambos contrastes se plantean de acuerdo a los supuestos establecidos en las hipótesis alternativas ( $H_1$ ) de los conjuntos de hipótesis 1 y 2.

Los estadísticos de prueba se calcularon a partir del siguiente modelo [2]:

$$t_{\Delta\bar{x}_{AF}} = \frac{\bar{x}_{AF,p} - \bar{x}_{AF,np}}{\sqrt{\frac{s_{AF,p}^2}{n_{AF,p}} + \frac{s_{AF,np}^2}{n_{AF,np}}}} \quad (4)$$

$$t_{\Delta \bar{x}_{DM}} = \frac{\bar{x}_{DM,p} - \bar{x}_{DM,np}}{\sqrt{\frac{s_{DM,p}^2}{n_{DM,p}} + \frac{s_{DM,np}^2}{n_{DM,np}}}} \quad (5)$$

Los elementos que integran las ecuaciones (4) y (5) se describen en la tabla 2:

Tabla 2. Descripción de los elementos para el cálculo de los estadísticos de prueba			
Métrica	Elemento	Descripción	Grupo
Anisotropía fraccional	$\bar{x}_{AF,p}$	Promedio	Portadores
	$s_{AF,p}^2$	Varianza	
	$n_{AF,p}$	Tamaño de la muestra	
	$\bar{x}_{AF,np}$	Promedio	No portadores
	$s_{AF,np}^2$	Varianza	
	$n_{AF,np}$	Tamaño de la muestra	
Difusividad media	$\bar{x}_{DM,p}$	Promedio	Portadores
	$s_{DM,p}^2$	Varianza	
	$n_{DM,p}$	Tamaño de la muestra	
	$\bar{x}_{DM,np}$	Promedio	No portadores
	$s_{DM,np}^2$	Varianza	
	$n_{DM,np}$	Tamaño de la muestra	

Los valores críticos se obtienen a partir de la distribución esperada para los estadísticos analizados que, en este caso, corresponden a las diferencias de medias para muestras recolectadas de poblaciones cuyas distribuciones se suponen normales o aproximadamente normales y para las cuales se desconocen las varianzas poblacionales y se suponen distintas. Esta última suposición se considera debido al desconocimiento de la verdadera distribución subyacente de las distribuciones y con el objetivo de integrar este factor de incertidumbre en las pruebas de hipótesis. Bajo estas consideraciones, se determina que la distribución que mejor describiría el comportamiento de las diferencias de medias muestrales es la distribución t-Student con  $v$  grados de libertad, en donde los grados de libertad para la distribución de la anisotropía fraccional se calculan con base en la siguiente expresión:

$$v_{AF} = \frac{\left( \frac{s_{AF,p}^2}{n_{AF,p}} + \frac{s_{AF,np}^2}{n_{AF,np}} \right)^2}{\frac{\left( \frac{s_{AF,p}^2}{n_{AF,p}} \right)^2}{n_{AF,p} - 1} + \frac{\left( \frac{s_{AF,np}^2}{n_{AF,np}} \right)^2}{n_{AF,np} - 1}} \quad (6)$$

Mientras que los grados de libertad para la distribución de la difusividad media se obtienen por:

$$v_{DM} = \frac{\left( \frac{s_{DM,p}^2}{n_{DM,p}} + \frac{s_{DM,np}^2}{n_{DM,np}} \right)^2}{\frac{\left( \frac{s_{DM,p}^2}{n_{DM,p}} \right)^2}{n_{DM,p} - 1} + \frac{\left( \frac{s_{DM,np}^2}{n_{DM,np}} \right)^2}{n_{DM,np} - 1}} \quad (7)$$

Finalmente, los valores críticos corresponden al valor de la variable aleatoria con distribución t-Student con  $v$  grados de libertad que dejan un área de  $\alpha$  del lado derecho o izquierdo de la distribución (dependiendo del tipo de contraste).

Una vez definido el modelo para la ejecución de las pruebas de hipótesis, se procedió a hacer el cálculo numérico a partir de los datos muestrales. Las diferencias para los promedios de las métricas entre los grupos de sujetos sanos portadores y no portadores se presentan en la Tabla 3 y se obtuvieron a partir del código presentado en la Figura 7.

Tabla 3. Promedios obtenidos para los grupos de portadores y no portadores de la mutación del gen			
Métrica	$\bar{x}_p$	$\bar{x}_{np}$	$\Delta \bar{x}$
Anisotropía Fraccional	0.3714	0.3842	-0.0128
Difusividad Media	$7.9283 \times 10^{-4}$	$7.7139 \times 10^{-4}$	$0.2144 \times 10^{-4}$

```
medias = struct('ad_adni',mean(ad_adni,1),'cnad_adni',mean(cnad_adni),...
mu_diff_dm = [(medias.portadores(1) - medias.no_portadores(1));...
(medias.no_portadores(1) - medias.ad_adni(1));...
(medias.portadores(1) - medias.cnad_adni(1));...
(medias.no_portadores(1) - medias.cnad_adni(1));...
(medias.portadores(1) - medias.dcl(1));...
(medias.no_portadores(1) - medias.dcl(1));...
(medias.portadores(1) - medias.cndcl(1));...
(medias.no_portadores(1) - medias.cndcl(1));...
(medias.portadores(1) - medias.no_portadores(1));...
(medias.ad_adni(1) - medias.cnad_adni(1));...
(medias.dcl(1) - medias.cndcl(1))]
mu_diff_af = [(medias.portadores(2) - medias.no_portadores(2));...
(medias.no_portadores(2) - medias.ad_adni(2));...
(medias.portadores(2) - medias.cnad_adni(2));...
(medias.no_portadores(2) - medias.cnad_adni(2));...
(medias.portadores(2) - medias.dcl(2));...
(medias.no_portadores(2) - medias.dcl(2));...
(medias.portadores(2) - medias.cndcl(2));...
(medias.no_portadores(2) - medias.cndcl(2));...
(medias.portadores(2) - medias.no_portadores(2));...
(medias.ad_adni(2) - medias.cnad_adni(2));...
(medias.dcl(2) - medias.cndcl(2))]
```

```
medias = struct with fields:
    ad_adni: [8.6458e-04 0.3321 0.0012 7.1298e-04 21.2245 77.1755]
    cnad_adni: [8.2398e-04 0.3445 0.0011 6.7170e-04 29.1333 74.1000]
    dcl: [8.2546e-04 0.3592 0.0011 6.6649e-04 26.7368 67.3158]
    cndcl: [8.1737e-04 0.3462 0.0011 6.6366e-04 29.1667 71.8095]
    no_portadores: [7.7139e-04 0.3842 0.0011 6.8694e-04 28.8750 36]
    portadores: [7.9283e-04 0.3714 0.0011 6.2890e-04 27.7714 36.0857]

mu_diff_dm = 11x1
10^-4 x
-0.9318
-0.3115
-0.5259
-0.3263
-0.5486
-0.2454
-0.4597
0.2144
0.4059
0.0809

mu_diff_af = 11x1
0.0528
0.0269
0.0397
0.0122
0.0249
0.0252
0.0388
```

Figura 7. Cálculo de los estadísticos muestrales necesarios para la ejecución de las pruebas.

Finalmente, a partir del cálculo de los elementos para las pruebas de hipótesis empleado los modelos discutidos previamente, y considerando un nivel de significancia del 5%, se obtienen las siguientes conclusiones:

Conjunto de hipótesis 1: *Existe evidencia estadísticamente suficiente para determinar que la anisotropía fraccional para el grupo de sujetos sanos portadores de la mutación E280A es significativamente menor que la del grupo de sujetos sanos no portadores de la mutación.*

Conjunto de hipótesis 2: *Existe evidencia estadísticamente suficiente para determinar que la difusividad media para el grupo de sujetos sanos portadores de la mutación E280A es significativamente mayor que la del grupo de sujetos sanos no portadores de la mutación.*

Los resultados de las pruebas se presentan en las figuras 8 y 9.

**Prueba de hipótesis para la diferencia de medias de la anisotropía fraccional para los portadores y los no portadores.**

```
n_p = length(portadores(:,2));
n_np = length(no_portadores(:,2));
var_p = var(portadores(:,2));
var_np = var(no_portadores(:,2));

v = ((var_np/n_np)+(var_p/n_p))^2/((var_np/n_np)^2/(n_np - 1) + (var_
t_val = tinv(0.95,v);
t_test = mu_dif_af(9,1)/sqrt(var_np/n_np + var_p/n_p);
if (t_test < t_val)
    disp("Existe evidencia estadísticamente suficiente para determina
    disp("que la anisotropía fraccional de los portadores es signific
    disp("menor que la de los no portadores")
else
    disp("No existe evidencia estadísticamente suficiente para deter
    disp("que la anisotropía fraccional de los portadores es signific
    disp("menor que la de los no portadores")
end
```

Existe evidencia estadísticamente suficiente para determinar  
que la anisotropía fraccional de los portadores es significativamente  
menor que la de los no portadores

Figura 8. Prueba de hipótesis para la anisotropía fraccional.

**Prueba de hipótesis para la diferencia de medias de la difusividad media para los portadores y los no portadores.**

```
var_p = var(portadores(:,1));
var_np = var(no_portadores(:,1));

v = ((var_np/n_np)+(var_p/n_p))^2/((var_np/n_np)^2/(n_np - 1) + (var_
t_val = tinv(0.05,v);

t_test = mu_dif_dm(9,1)/sqrt(var_np/n_np + var_p/n_p);

if (t_test > t_val)
    disp("Existe evidencia estadísticamente suficiente para determina
    disp("que la difusividad media de los portadores es significativa
    disp("mayor que la de los no portadores")
else
    disp("No existe evidencia estadísticamente suficiente para determ
    disp("que la difusividad media de los portadores es significativa
    disp("mayor que la de los no portadores")
end
```

Existe evidencia estadísticamente suficiente para determinar  
que la difusividad media de los portadores es significativamente  
mayor que la de los no portadores

Figura 9. Prueba de hipótesis para la difusividad media.

## Referencias

[1] Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association, 46(253), 68–78. <https://doi.org/10.2307/2280095>

[2] Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers y Keying Ye (2012). Probabilidad y Estadística para Ingeniería y Ciencias, 9na. Edición. Person Education, México.