TEMA 0. REPASO BÁSICO DE CONCEPTOS DE ESTADÍSTICA

ÍNDICE

TEMA 0. REPASO BÁSICO DE CONCEPTOS DE ESTADÍSTICA

La estadística es la rama de las matemáticas que se encarga del estudio de determinadas características en una población, recogiendo los datos, agrupándolos, organizándolos en tablas, representándolos gráficamente y analizándolos para sacar conclusiones de dicha población.

Podemos considerar dos grandes áreas: la estadística descriptiva y la estadística inferencial. Esta última es la que se tratará en el temario correspondiente a la asignatura de Diseños de investigación y análisis de datos. Por lo tanto, en este tema 0 nos centraremos en un repaso rápido de la descriptiva.

0.1. Estadística descriptiva

0.1.1. Conceptos iniciales

Se organizan y se resumen conjuntos de observaciones cuantificadas procedentes de una muestra o de la población total.

Población: Conjunto de todos los elementos que cumplen una determinada característica objeto de estudio.

Muestra: Es un subconjunto cualquiera de una población.

Parámetro: Es una propiedad descriptiva (una medida) de una población, Se denota con letras griegas.

Estadístico: Es una propiedad descriptiva (una medida) de una muestra. Se denota con letras latinas.

	MUESTRA	POBLACIÓN
MEDIA	$\bar{X}$	μ
VARIANZA SESGADA	$S_n^2 = S_X^2$	σ_x^2
CUASIVARIANZA	$S_{n-1}^2 = \hat{S}_X^2$	
DESVIACIÓN TÍPICA SESGADA	$S_n = S_x$	σ_x
CUASIDESVIACIÓN TÍPICA	$S_{n-1} = \hat{S}_x$	
PROPORCIÓN	p	π

En los próximos apartados se verá qué es cada una de estas medidas.

VARIABLES

Una variable es un conjunto de valores resultantes de medir una característica de interés sobre cada elemento individual de una población o muestra. Se representan con letras latinas en mayúsculas.

Aquí tenéis una tabla resumen con todos los tipos de variables y sus escalas de medida.

Tipo de variable	Escala de medida	Características básicas	Relaciones válidas	Ejemplos
CUALITATIVA: - Dicotómica - Politómica	Nominal	Los números identifican y clasifican objetos	Relaciones del tipo "igual que" o "distinto que"	Sexo, estado civil, raza, diagnóstico clínico
CUASICUANTITATIVA	Ordinal	Además, los números indican las posiciones relativas de los objetos	Además, relaciones del tipo "mayor que" o "menor que"	Dureza, posición en el ranking ATP, grado de satisfacción.
CUANTITATIVA: - Discreta - Continua	Intervalo	Además, hay una unidad de medición común	Además, igualdad o desigualdad de diferencias	Temperatura en grados centígrados, inteligencia
	Razón	Además, el punto cero es absoluto	Además, igualdad o desigualdad de razones	Longitud, peso, altura, tiempo de reacción

0.1.2. Índices de tendencia central.

Dentro de los índices de tendencia central, tenemos la moda, la mediana y la media. En este repaso solo vamos a hacer hincapié en la media.

MEDIA

Se define como la suma de todas las puntuaciones de la distribución, dividida por el total de los casos u observaciones.

$$\bar{X} = \frac{\sum n_i X_i}{n}$$

donde:

n = Número total de observaciones

X_i = Cada uno de los valores de la variable o el punto medio si tenemos intervalos

n_i = Frecuencia absoluta de valor X o del intervalo

Vamos a calcular la media con la distribución de frecuencias de los alumnos que han aprobado el examen de Introducción al Análisis de datos de la UNED.

X	n_i	$n_i \cdot X_i$
5	5	25
6	7	42
7	8	56
8	7	56
9	1	9
10	1	10
	n=29	198

$$\bar{X} = \frac{\sum n_i X_i}{n} = \frac{198}{29} = 6{,}83$$

0.1.3. Índices de posición

Son los índices que permiten establecer la posición relativa de una puntuación respecto al grupo de puntuaciones y se suelen llamar cuantiles.

En este repaso lo que nos interesa es entender qué información nos facilita este tipo de índices, el cálculo como tal no lo necesitamos de cara a la asignatura de Diseños de investigación.

Dentro de los cuantiles podemos tener: percentiles, cuartiles y deciles.

PERCENTILES

Son los 99 valores que dividen en 100 partes iguales la distribución de frecuencias de la variable.

CUARTILES

Son los 3 valores que dividen en 4 partes secciones, cada una conteniendo el 25% de las observaciones.

DECILES

Son los 9 valores que dividen en 10 partes iguales la distribución de frecuencias de la variable.

0.1.4. Medidas de variabilidad

Las medidas de variabilidad nos informan sobre en qué medida las puntuaciones en una determinada variable están próximas o alejadas del índice de tendencia central.

En el temario correspondiente a Introducción al análisis de datos se ven varias medidas de variabilidad, pero en este repaso nos centraremos en cuatro de ellas, que son con las que trabajaremos a lo largo del temario de Diseños de investigación.

VARIANZA SESGADA

Se representa por S^2_x y se define como el promedio de los cuadrados de las desviaciones de las puntuaciones con respecto a la media. Siempre es positiva.

$$S_X^2 = \frac{\sum n_i \cdot X_i^2}{n} - \bar{X}^2$$

DESVIACIÓN TÍPICA SESGADA

Se representa por S_x y es la raíz cuadrada positiva de la varianza.

$$S_X = \sqrt{S_X^2}$$

CUASIVARIANZA

Se representa por S^2_{n-1} y también se la conoce como varianza insesgada. Se calcula a través de la siguiente fórmula.

$$S_{n-1}^2 = \frac{\sum n_i (X_i - \bar{X})^2}{n-1}$$

CUASIDESVIACIÓN TÍPICA

Se representa por S_{n-1} y también se la conoce como desviación típica insesgada. Se calcula a través de la siguiente fórmula.

$$S_{n-1} = \sqrt{\frac{\sum n_i (X_i - \bar{X})^2}{n-1}}$$

Vamos a calcularlas con la distribución de frecuencias de los alumnos que han aprobado el examen de Introducción al Análisis de datos de la UNED.

X	n_i	$n_i \cdot X^2$	$n_i \cdot (X-\bar{X})^2$
5	5	125	16,74
6	7	252	4,82
7	8	392	0,23
8	7	448	9,58
9	1	81	4,71
10	1	100	10,05
	n=29	1398	46,13

$$\bar{X} = 6{,}83$$

$$S_X^2 = \frac{\sum n_i \cdot X_i^2}{n} - \bar{X}^2 = \frac{1398}{29} - 6{,}83^2 = 1{,}58$$

$$S_X = \sqrt{S_X^2} = \sqrt{1{,}58} = 1{,}26$$

$$S_{n-1}^2 = \frac{\sum n_i (X_i - \bar{X})^2}{n-1} = \frac{46{,}13}{29-1} = 1{,}65$$

$$S_{n-1} = \sqrt{S_{n-1}^2} = \sqrt{1{,}65}$$

0.1.5. Puntuaciones directas, diferencias y típicas

Las puntuaciones directas son con las que tratamos habitualmente (puntuaciones de un sujeto en un test, etc.), pero la comparación de las puntuaciones directas de un mismo sujeto en dos variables distintas puede llevarnos a confusión, ya que las puntuaciones directas nos ofrecen muy poca información. De hecho, conocida una puntuación directa no sabemos si se trata de un valor alto o bajo porque esto depende del promedio del grupo.

Una posible solución es trabajar con **puntuaciones diferenciales.** Si a una puntuación directa le restamos la media de su grupo obtenemos una puntuación diferencial.

$$x_i = X_i - \bar{X}$$

Aportan más información ya que podemos saber si la puntuación coincide con la media, si es inferior o es superior a ella.
Estas puntuaciones tienen dos propiedades:

- La media es 0
- La varianza de las puntuaciones diferenciales es igual a la varianza de las puntuaciones directas.

Sin embargo, dos puntuaciones diferenciales idénticas pueden tener un significado muy diferente en función de la media y de la varianza de las distribuciones de las que proceden. Para eliminar este inconveniente se utilizan las puntuaciones típicas. Las puntuaciones típicas van más allá y nos permiten no sólo comparar las puntuaciones de un sujeto en dos variables distintas, sino también comparar dos sujetos distintos en dos pruebas o variables distintas.

Una puntuación típica se calcula con la siguiente fórmula:

$$z_x = \frac{x_i}{S_X} = \frac{X_i - \bar{X}}{S_X}$$

Al proceso de obtener las puntuaciones típicas se llama tipificación. Una puntuación típica indica el número de desviaciones típicas que se aparta de la media una determinada puntuación.
Sus propiedades son:

- La media es igual a 0.
- La varianza es igual a 1

0.1.6. Relación entre variables

Este apartado y el modelo de regresión se desarrollarán en el temario propio de la asignatura de Diseños de investigación.

0.2. Modelos continuos de probabilidad

Vamos a ver los modelos de distribución para variables aleatorias continuas más utilizados en el área de Psicología y Ciencias de la Salud. Conviene distinguir entre aquellas distribuciones de probabilidad a las que frecuentemente se ajustan las variables con las que trabajamos (distribución normal) y aquellas distribuciones que tienen una gran aplicación como instrumentos estadísticos (distribución χ^2 de Pearson, la t de Studente y la F de Fisher-Snedecor).

Estas tres últimas derivan de la distribución normal y tienen una gran importancia como instrumentos estadísticos en la estadística inferencial.

0.2.1. La distribución normal

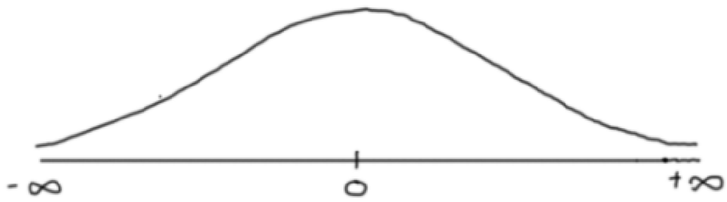
También llamada Campana de Gauss o curva normal.

Su figura nos indica que la puntuación de la mayoría de los individuos es una variable que sigue esta distribución, se encuentra en torno a la media y, a medida que nos alejamos de esa puntuación, por su lado izquierdo y derecho, va disminuyendo la frecuencia.

Si restamos la media y dividimos por la desviación típica obtenemos una nueva variable que designamos por z.

$$z = \frac{X - \mu}{\sigma}$$

Esta nueva variable z se distribuirá normalmente con media igual a cero y desviación típica igual a 1. Esta distribución se denomina **normal tipificada o normal estandarizada**. Para la aplicación a problemas concretos en que se siga esta distribución recurriremos a las tablas correspondientes del formulario.



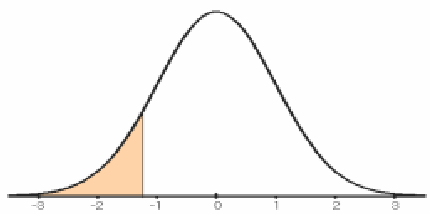
UTILIZACIÓN DE LAS TABLAS

En las Tablas III y IV se recoge la función de distribución de la curva normal estándar. En ellas se presentan todas las puntuaciones típicas desde -3,59 hasta +3,59 con intervalos de 0,01.

La primera columna, encabezada con la letra z, consta de un número con un decimal, que corresponde a la puntuación típica. Y la primera fila (a la derecha de la z) corresponde al segundo decimal de la puntuación z. Todos los valores interiores representan probabilidades. La Tabla III corresponde a las puntuaciones negativas (por debajo de la media) y la Tabla IV a las positivas (por encima de la media).

Por ejemplo, la puntuación típica z = -2,97 (Tabla III) deja por debajo de sí una probabilidad de 0,0015.

Tabla III: Distribución NORMAL TIPIFICADA



$P(Z \leq z)$

z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,50	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002
-3,40	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002
-3,30	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003
-3,20	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005
-3,10	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
-3,00	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2,90	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	$\boxed{0,0015}$	0,0014	0,0014
-2,80	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,70	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,60	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,50	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,40	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,30	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,20	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,10	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,00	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,90	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

La puntuación típica z = 2,97 (Tabla IV) deja por debajo de sí una probabilidad de 0,9985.

Al ser una distribución simétrica puede comprobarse que la proporción que queda por debajo de z = -2,97 es igual a la proporción que queda por encima de z = 2,97.

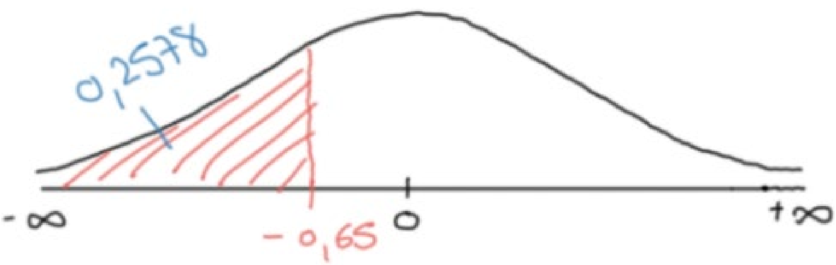
→ Cálculo de la probabilidad para valores menores o iguales que una determinada puntuación típica.

En este caso se busca directamente en la tabla.

Si una variable se distribuye normalmente, ¿cuál es la probabilidad de obtener valores menores o iguales que z = - 0,65?

Como el valor es negativo lo buscamos en la Tabla III, buscamos en la primera columna el -0,60 y en la primera fila el 0,05.

z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,50	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002
-3,40	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002
-3,30	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003
-3,20	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005
-3,10	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
-3,00	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-0,90	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,80	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,70	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,60	0,2743	0,2709	0,2676	0,2643	0,2611	$\boxed{0,2578}$	0,2546	0,2514	0,2483	0,2451
-0,50	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,40	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121



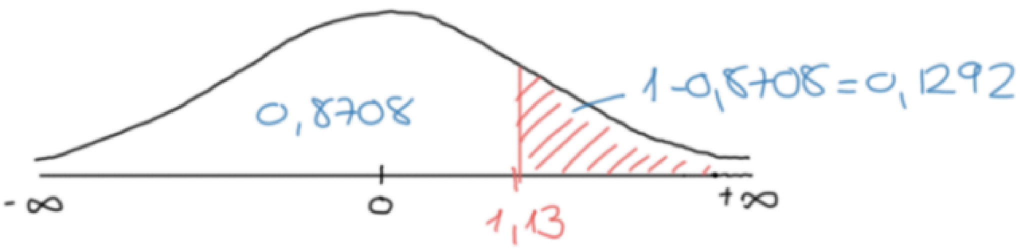
→ Cálculo de la probabilidad para valores mayores que una determinada puntuación típica.

En este caso se mira en la tabla la probabilidad que esa puntuación deja por debajo y se resta de 1.

Si una variable se distribuye normalmente, ¿cuál es la probabilidad de obtener valores mayores que z = 1,13?

Si se mira la Tabla IV, la puntuación típica 1,13 deja por debajo de sí una probabilidad de 0,8708. Como lo que se pregunta es por la probabilidad que queda por encima, para calcularla restaremos esa probabilidad de 1.

1 – 0,8708 = 0,1292



→ Otros casos interesantes en el uso de la tabla de la curva normal

Vamos a ver un par de ejemplos.

Las puntuaciones de una determinada asignatura, X, de un grupo de 500 niños se distribuyen normalmente con media 6 y desviación típica 2. ¿Cuántos niños no han alcanzado la puntuación 5?

Lo primero es transformar la puntuación directa, X = 5 en una puntuación típica para poder buscarla en tablas:

$$z = \frac{5-6}{2} = -0{,}5$$

En la Tabla III se observa que esta puntuación deja por debajo de sí una proporción de 0,3085, que es el 30,85%.

Por lo tanto, 0,3085 · 500 = 154,25 → 154 niños

Siguiendo con los mismos datos, ¿cuál será el percentil 75 de la distribución?

Ya sabemos que el percentil 75 es una puntuación directa que deja por debajo de sí el 75% de los casos. A este percentil le corresponde una puntuación típica que deja por debajo de sí una proporción de casos de 0,75.

En este caso entramos en la Tabla IV (probabilidades por encima de 0,5) y buscamos en el interior la probabilidad de 0,75 o la más cercana y así obtenemos la puntuación típica correspondiente.

z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,00	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,10	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,20	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,30	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,40	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,50	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,60	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	$\boxed{0,7486}$	0,7517	0,7549
0,70	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852

Una vez que tenemos la puntuación típica obtenemos la puntuación directa correspondiente al percentil 75 de la siguiente manera:

$$z = \frac{X - \bar{X}}{S_x} \qquad 0,67 = \frac{X - 6}{2} \qquad 0,67 \cdot 2 = X - 6 \qquad X = 0,67 \cdot 2 + 6 = 7,34$$

Por lo tanto, el percentil 75 corresponde a una puntuación directa de 7,34. Quiere decir que hay un 75% de niños con una puntuación menor de 7,34 y un 25% con una puntuación mayor de 7,34.

0.2.2. La distribución t de Student

Al igual que la distribución normal es simétrica con media cero. Su forma es muy parecida a la normal tipificada pero menos apuntada. Para efectos prácticos, nosotros las dibujaremos iguales.

En estas aparece un nuevo concepto que son los grados de libertad (n) que son necesarios para entrar en las tablas correspondientes a la t de Student.

En la Tabla VI se presentan los valores positivos para esta distribución. En la primera columna se presentan los grados de libertad y en la primera fila las distintas probabilidades o proporciones de valores menores o iguales que un valor positivo dado. Como se trata de una distribución simétrica podemos hallar las probabilidades asociadas a los valores negativos a partir de los positivos.

Vamos a verlo con un ejemplo:

Sea X una variable que se distribuye según t con 5 grados de libertad.

a) ¿Cuál es la probabilidad de obtener valores menores o iguales a 2,015?

Consultamos la Tabla VI. En la primera columna (grados de libertad) localizamos el valor 5. Los valores incluidos en su fila correspondiente son valores de t. Localizamos 2,015 y se ve que en la primera fila corresponde con 0,95. Por lo tanto:

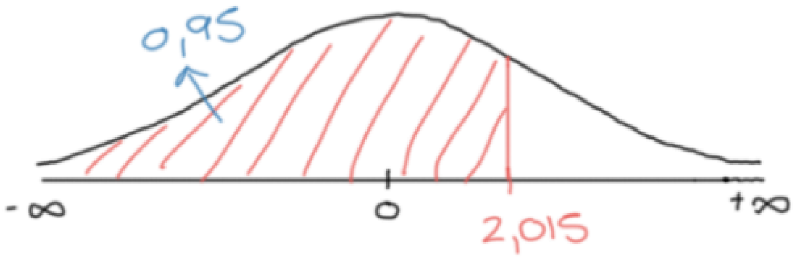
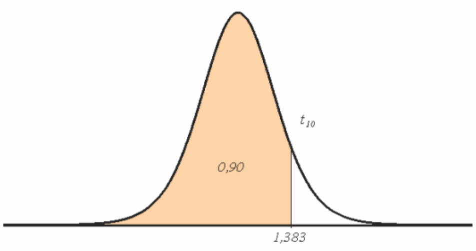


Tabla VI: Distribución t de Student



$P(T \leq t_{gl})$

	Probabilidad											
g.l.	**0,550**	**0,600**	**0,650**	**0,700**	**0,750**	**0,800**	**0,850**	**0,900**	**0,950**	**0,975**	**0,990**	**0,995**
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707

b) Calcular P (X > 0,920)

En la Tabla VI vemos que para t_5: P(X≤0,920) = 0,80

Por lo tanto, P(X>0,920) = 1-0,80 = 0,20

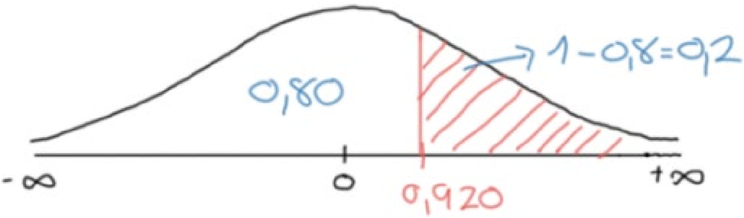
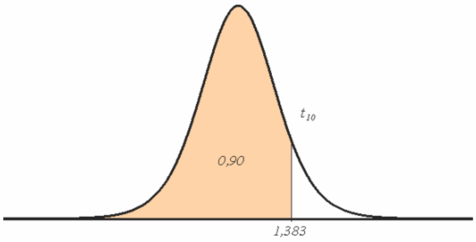


Tabla VI: Distribución t de Student



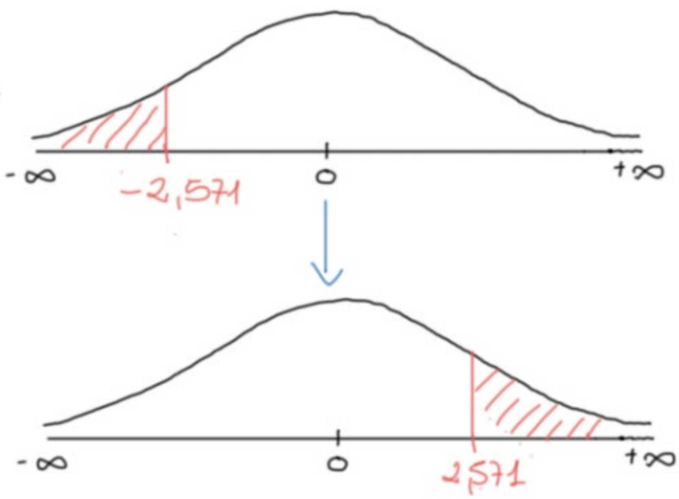
$P(T \leq t_{gl})$

g.l.	Probabilidad 0,550	0,600	0,650	0,700	0,750	0,800	0,850	0,900	0,950	0,975	0,990	0,995
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707

c) Calcular P (X ≤ -2,571)

En el formulario solo tenemos las tablas de las puntuaciones positivas de la t de Student, es decir las que probabilidades por encima de 0,5. Pero sabemos que la distribución t de Student es simétrica, por lo tanto:

P (X ≤ -2,571) = P (X > 2,571) = 1 - P (X ≤ 2,571) = 1 – 0,975 = 0,025



0.2.3. La distribución χ^2 de Pearson

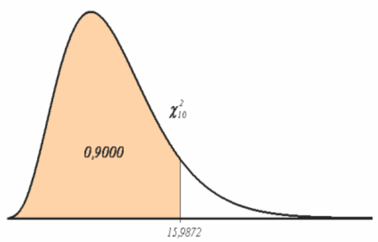
Una de las propiedades que tiene esta distribución es que no adoptan valores menores de 0. Es asimétrica positiva.

La Tabla V permite obtener las probabilidades acumuladas a algunos valores de esta distribución. La primera fila recoge las probabilidades o proporciones y la primera columna los grados de libertad correspondientes. En el interior de la tabla se encuentran los valores de la variable.

Tenemos una variable que sigue una distribución χ^2 con 5 grados de libertad. ¿Qué proporción deja por debajo el valor de 11,07?

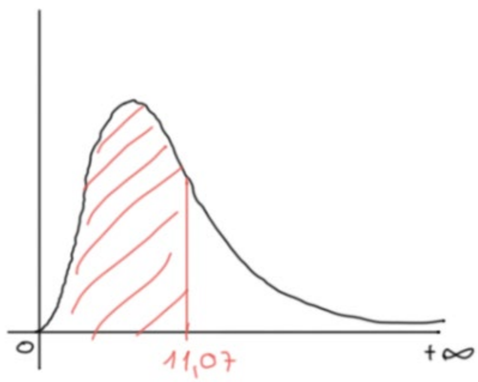
$P(X \leq 11{,}07) = 0{,}95$

Tabla V: Distribución CHI-CUADRADO



$$P\left(X \leq \chi^2_{gl}\right)$$

g.l.	Probabilidad									
	0,005	0,010	0,025	0,050	0,100	0,900	0,950	0,975	0,990	0,995
1	0,0000	0,0002	0,0010	0,0039	0,0158	2,7055	3,8415	5,0239	6,6349	7,8794
2	0,0100	0,0201	0,0506	0,1026	0,2107	4,6052	5,9915	7,3778	9,2103	10,5966
3	0,0717	0,1148	0,2158	0,3518	0,5844	6,2514	7,8147	9,3484	11,3449	12,8382
4	0,2070	0,2971	0,4844	0,7107	1,0636	7,7794	9,4877	11,1433	13,2767	14,8603
5	0,4117	0,5543	0,8312	1,1455	1,6103	9,2364	11,0705	12,8325	15,0863	16,7496
6	0,6757	0,8721	1,2373	1,6354	2,2041	10,6446	12,5916	14,4494	16,8119	18,5476
7	0,9893	1,2390	1,6899	2,1673	2,8331	12,0170	14,0671	16,0128	18,4753	20,2777



0.2.4. La distribución F de Fisher-Snedecor

Es una distribución asimétrica positiva, por lo que nunca toma valores inferiores a 0.

La Tabla VII recoge solamente la probabilidad de que X sea menor o igual que 0,900; 0,950; 0,975; 0,990 y 0,995, que son los valores utilizados habitualmente.

Sea una variable X que se distribuye según la distribución F de Fisher-Snedecor con 5 y 10 grados de libertad.

a) P(X≤3,326)

Vamos entrando con 5 grados de libertad en el numerador (columna) y 10 para el denominador (fila) en cada una de las tablas buscando el valor de 3,326. En este caso corresponde a 0,95

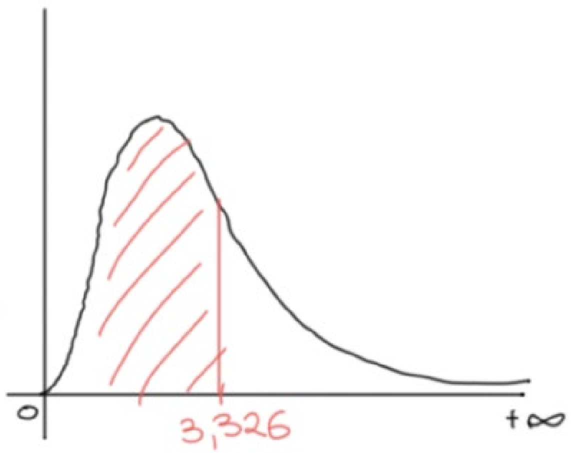
Tabla VII: Distribución F de Snedecor-Fisher

$$P\left(F_{n_1,n_2} \leq f_{n_1,n_2}\right) = 0{,}95$$

Grados de libertad del numerador (n_1)

n_2	1	2	3	4	5	6	7	8	9	10	20	30	40	50	60	120
1	161,448	199,500	215,707	224,583	230,162	233,986	236,768	238,883	240,543	241,882	248,013	250,095	251,143	251,774	252,196	253,253
2	18,513	19,000	19,164	19,247	19,296	19,330	19,353	19,371	19,385	19,396	19,446	19,462	19,471	19,476	19,479	19,487
3	10,128	9,552	9,277	9,117	9,013	8,941	8,887	8,845	8,812	8,786	8,660	8,617	8,594	8,581	8,572	8,549
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999	5,964	5,803	5,746	5,717	5,699	5,688	5,658
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,772	4,735	4,558	4,496	4,464	4,444	4,431	4,398
6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147	4,099	4,060	3,874	3,808	3,774	3,754	3,740	3,705
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,677	3,637	3,445	3,376	3,340	3,319	3,304	3,267
8	5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438	3,388	3,347	3,150	3,079	3,043	3,020	3,005	2,967
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,179	3,137	2,936	2,864	2,826	2,803	2,787	2,748
10	4,965	4,103	3,708	3,478	$\boxed{3{,}326}$	3,217	3,135	3,072	3,020	2,978	2,774	2,700	2,661	2,637	2,621	2,580
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,896	2,854	2,646	2,570	2,531	2,507	2,490	2,448
12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,796	2,753	2,544	2,466	2,426	2,401	2,384	2,341
13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,714	2,671	2,459	2,380	2,339	2,314	2,297	2,252

La puntuación de 3,326 corresponde con el percentil 95.



b) Determinar el valor del percentil 5.

En este caso tenemos que aplicar la propiedad recíproca. Buscamos en la misma tabla que antes, en la de 0,95, pero girando los grados de libertad. Y hacemos 1 partido ese valor.

Tabla VII: Distribución F de Snedecor-Fisher

$$P\left(F_{n_1,n_2} \le f_{n_1,n_2}\right) = 0{,}95$$

Grados de libertad del numerador (n_1)

n_2	1	2	3	4	5	6	7	8	9	10	20	30	40	50	60	120
1	161,448	199,500	215,707	224,583	230,162	233,986	236,768	238,883	240,543	241,882	248,013	250,095	251,143	251,774	252,196	253,253
2	18,513	19,000	19,164	19,247	19,296	19,330	19,353	19,371	19,385	19,396	19,446	19,462	19,471	19,476	19,479	19,487
3	10,128	9,552	9,277	9,117	9,013	8,941	8,887	8,845	8,812	8,786	8,660	8,617	8,594	8,581	8,572	8,549
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999	5,964	5,803	5,746	5,717	5,699	5,688	5,658
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,772	$\boxed{4{,}735}$	4,558	4,496	4,464	4,444	4,431	4,398
6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147	4,099	4,060	3,874	3,808	3,774	3,754	3,740	3,705
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,677	3,637	3,445	3,376	3,340	3,319	3,304	3,267
8	5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438	3,388	3,347	3,150	3,079	3,043	3,020	3,005	2,967
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,179	3,137	2,936	2,864	2,826	2,803	2,787	2,748
10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	3,020	2,978	2,774	2,700	2,661	2,637	2,621	2,580
11																2,448
12																2,341
13																2,252

Por lo tanto, será: $\frac{1}{4{,}735} = 0{,}211$