TEMA 1

INFERENCIA ESTADÍSTICA Y DISTRIBUCIONES EN EL MUESTREO

Introducción al tema 1

Hemos dicho que una sucesión de variables aleatorias en un espacio muestral Ω es un conjunto de variables aleatorias $\{x_1, x_2, \dots, x_n, \dots\}$, conjunto que ha de ser o bien finito o bien numerable. Es decir que cada X_i es una función en la forma:

$$\begin{aligned} X_i : \Omega &\longrightarrow \mathbb{R} \\ w &\longrightarrow x_i(w) \end{aligned}$$

Esta definición es muy abstracta, pongamos un ejemplo:

Sea $\{X_i\}_{1 \le i \le 365}$ la temperatura que hizo en Murcia el día que indica el subíndice. El espacio muestral Ω es el conjunto de temperaturas diarias en Murcia (presentes, pasadas o futuras, en cualquier año) El espacio Ω es muy grande y desconocemos la mayoría de sus datos (por ejemplo las temperaturas en el siglo XII seguro que no se conocen)

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ w &\longrightarrow X(w) = X_w \end{aligned}$$

Así si w = temperatura el día 3 de enero, entonces $X(w) = X_w = X_3$

La distribución de Poisson se emplea para describir sucesos que se repiten en el tiempo. No estoy seguro, pero tal vez las temperaturas diarias siguen una distribución de Poisson ya que se repiten en el mismo día del año, quiero decir, que la temperatura del 17 de marzo de 2010 es similar a la del 17 de marzo de 2011, esta es la razón de que piense en la distribución de Poisson

De esta manera $X_i \rightsquigarrow P(\lambda_i)$, pero no sabemos el valor de λ_i, aquí está el problema. Una manera de estimar α_i sería hallar la media de las temperaturas registradas en el día i. Por ejemplo si disponemos de datos desde 1945 a estas fechas,

$$\alpha_1 = \frac{\sum \text{temperaturas el día 1 de enero}}{66}$$

Lo mismo podríamos hacer con el resto de los parámetros

Volvamos a nuestra sucesión $X_1, X_2, \dots X_n, \dots$

¿Qué sería una muestra de tamaño 4?

Pues sería uno de los siguientes conjuntos

- $\{X_1, X_{200}, X_{302}, X_{350}\}$
- $\{X_3, X_{150}, X_{64}, X_{292}\}$
- $\{X_1, X_{45}, X_{302}, X_3\}$

¿Qué sería una realización muestral para la muestra de tamaño 4 $\{X_1, X_2, X_{302}, X_{350}\}$?

Pues sería el valor que toman las temperaturas en esos días para un año determinado. Por ejemplo si en 2004 $X_1 = 2°C$, $X_{200} = 38°C$, $X_{302} = 3°C$ $X_{350} = -1°C$, una realización muestral para la muestra de tamaño 4 $\{X_1, X_{200}, X_{302}, X_{350}\}$ sería $\{2, 38, 3, -1\}$

Nos preguntamos si las muestras de tamaño n anteriormente mencionadas son m.a.s. La respuesta es que no.

Si tomamos una muestra de tamaño n $X_1, X_2, \dots X_n$ es evidente que las variables X_i son independientes (la temperatura de un día no depende de la que hizo o vaya a hacer en otro día cualquiera). Las X_i no están idénticamente distribuidas ya que si bien todas se distribuyen según una distribución de Poisson, el parámetro es distinto para cada variable. λ mide la temperatura media en un día determinado, así por ejemplo $X_1 \rightsquigarrow P(10)$, $X_{200} \rightsquigarrow P(32)$. Por tanto las muestras anteriores no son m.a.s.

Diferencia entre media muestral y media poblacional

Sea $X_1, X_2, \dots X_n$ una m.a.s. en una población que tiene como función de probabilidad $P(X=x)$ si se trata de variable discreta, o bien con función de densidad $f(x)$ si se trata de una variable continua.

Se define la media muestral como:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Se define la media poblacional como:

$$E[X] = \mu = \sum_{i=1}^{\infty} X_i \cdot P(X = X_i) \quad (\text{variable discreta})$$

$$E[X] = \mu = \int_{-\infty}^{\infty} x\, f(x)\, dx \quad (\text{variable continua})$$

Para hallar la media poblacional se necesita saber la función de distribución de X, cosa que en la mayoría de los casos es desconocida. Esto se puede aplicar a cualquier parámetro de la población. Se define la varianza muestral como:

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}$$

La varianza poblacional valdría

$$\sigma^2 = Var[X] = E[(X-\mu)^2] = \sum_{i=1}^{\infty} (X_i - \mu)^2 \cdot P[X = X_i]$$

o bien:

$$\sigma^2 = Var[X] = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\, dx$$

TEMA 1. INFERENCIA ESTADÍSTICA Y DISTRIBUCIONES EN EL MUESTREO

Supongamos que quisiéramos hacer una predicción sobre los resultados electorales en las próximas elecciones. Hay un método infalible para llevar a cabo con éxito esto, si ningún ciudadano con derecho a voto miente y le preguntamos a todos y cada uno, el resultado de la encuesta sería al 100% fiable. Pero esto no es posible, a parte de no tener el más mínimo control para saber si los encuestados mienten, sería muy costoso y llevaría mucho tiempo el realizarlo. ¿Qué hacemos? Procederíamos a estudiar una muestra, más o menos grande, que pudiera representar a la población.

A partir de esta muestra, trasladaríamos las conclusiones a la población. Concluiríamos hechos de un caso particular a un caso general, el método se llama método inductivo, al contrario de lo que hace la matemática que utiliza el método deductivo, es decir aplicar las consecuencias de un caso general a un caso particular. La inferencia estadística consiste en esto, en estudiar la población a partir de muestras, aplicando inducción. Esto conlleva un riesgo, como todo método inductivo, ir de casos particulares a casos generales,

puede provocar la situación que se produjo en un pequeño pavo que llegó muy temeroso a una granja el día 1 de enero. Por cierto, ese mismo día a las 9h. le dieron de comer, al igual que el día 2 a la misma hora y los días posteriores. Allá por el 20 de agosto dedujo una ley: "todos los días, a las 9h. de la mañana, me echarán de comer". El pavo inductivista se llevó una gran sorpresa el día 24 de diciembre, en vez de echarle de comer le cortaron el pescuezo. Hay que llevar mucho cuidado con los procesos inductivos.

El planteamiento del problema es el siguiente: una variable aleatoria que describe cierta característica de la población pero no conocemos con profundidad la población:

1º Conocemos la función de distribución de la variable X pero no conocemos algún parámetro θ. En este caso hacemos inferencia paramétrica. Por ejemplo si sabemos que X se distribuye normalmente, con varianza 9 y media desconocida $X \rightsquigarrow N(\mu, 3)$

2º No conocemos la función de distribución de X, $F(x)$. En este caso estamos haciendo estadística no paramétrica.

CONCEPTOS BASICOS

Población: conjunto de individuos que tienen una característica en común, que se mide a través de una variable aleatoria.

Muestra: es un subconjunto de elementos de una población, el tamaño de la muestra es el $\underbrace{\text{cardinal}}_{\text{nº elementos subconjunto}}$ de este subconjunto.

Realización muestral: es un conjunto concreto de valores de la muestra de tamaño n, $x_1, x_2, \dots x_n$.

Muestreo: es el proceso seguido para seleccionar la muestra. Distinguimos dos tipos de muestreos:

a) Muestreo con reemplazamiento: siempre estamos en las mismas condiciones puesto que todo elemento que sale al tomar la muestra se devuelve al tomar otra muestra.

b) Muestreo sin reemplazamiento: nunca estamos en las mismas condiciones, conforme vamos seleccionando sucesivamente las muestras hay menos individuos en la población.

Por ejemplo, cuando en una urna que contiene N bolas hacemos n-extracciones (es lo mismo que seleccionar una muestra de tamaño n):

a) Si la extracción es con reemplazamiento la probabilidad de extraer una bola en concreto es:

$$p = \frac{1}{N} \cdot \frac{1}{N} \overset{n \text{ veces}}{\cdots\cdots} \frac{1}{N} = \left(\frac{1}{N}\right)^n$$

b) Si la extracción es sin reemplazamiento la probabilidad de extraer una bola determinada es:

$$P = \frac{1}{N} \cdot \frac{1}{N-1} \cdot \frac{1}{N-2} \cdots \frac{1}{N-(n-1)}$$

Para poblaciones infinitas ($N \longrightarrow \infty$) los dos muestreos ofrecen los mismos resultados.

MUESTREO ALEATORIO SIMPLE (m.a.s.)

La selección de la muestra tiene como fin el estudio de la población. Conviene que la muestra seleccionada sea representativa de la población. El muestreo aleatorio simple consiste en seleccionar una muestra teniendo en cuenta las siguientes hipótesis:

1º La muestra se obtiene al azar de modo que todos los individuos tienen la misma probabilidad de ser seleccionados.

2º La función de distribución de la población es válida para cualquier muestra de la misma

3º Las variables muestrales $x_1, x_2, \dots x_n$ son independientes.

Observación

* Recordemos la definición de probabilidad condicionada: La probabilidad de que se verifique A una vez que se ha verificado B se representa por $P(A/B)$ y se define mediante la expresión:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

* Recordemos la definición de sucesos independientes
Dos sucesos A y B se dicen independientes si verifican que:
$P(A/B) = P(A)$; $P(B/A) = P(B)$.

Proposición: Sean A y B dos sucesos independientes. Entonces:

$P(A \cap B) = P(A) \cdot P(B)$

Demostración:

Al ser A y B independientes $P(A/B) = P(A) \Rightarrow \dfrac{P(A \cap B)}{P(B)} = P(A) \Rightarrow$

$\Rightarrow P(A \cap B) = P(A) \cdot P(B)$

Función de distribución de una m.a.s.

Sea $\{X_1, X_2, \dots, X_n\}$ una muestra aleatoria simple. Supongamos que la función de distribución de X_i es $F_{X_i}(x_i)$. Entonces como las variables X_i son independientes, si $F(x_1, x_2, \dots x_n)$ es la función de distribución de la muestra aleatoria simple:

$$F(x_1, x_2, \dots x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) =$$

$$= P(X_1 \leq x_1 \cap X_2 \leq x_2 \cap \dots \cap X_n \leq x_n) = P(X_1 \leq x_1) \cdot P(X_2 \leq x_2) \cdot \dots$$

$$\dots P(X_n \leq x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \dots F_{X_n}(x_n) = \prod_{i=1}^{n} F_X(x_i)$$

y esta igualdad ultima igualdad se verifica porque al estar las X_i identicamente distribuidas la función de distribución es la misma para todas, es decir:

$F_{X_1}(x_1) = F_X(x_1)$; $F_{X_2}(x_2) = F_X(x_2) \dots F_{X_n}(x_n) = F_X(x_n)$

Ejemplo:

En una urna hay 2 bolas blancas, una amarilla y tres rojas. Extraemos una bola y definimos la siguiente variable:

X (Blanca) $= 0$; X (amarilla) $= 1$; X (roja) $= 2$

a) Calcular la función de probabilidad de X

x	0	1	2
$P(X=x)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

b) Extraemos una m.a.s. de tamaño 2 de la población anterior. Obtener la función de probabilidad de la muestra aleatoria simple.

Si $\{X_1, X_2\}$ es una muestra aleatoria simple, X_1 y X_2 tienen la misma función de probabilidad ya que X_1 y X_2 son independientes por lo siguiente:

X_1 es el resultado de la primera extracción, miramos el resultado, lo anotamos devolviendo la bola a la urna y acto seguido se procede a la segunda extracción.

Si nos fijamos una m.a.s. de tamaño 2 se puede interpretar como una variable aleatoria bidimensional (X_1, X_2).
En general una m.a.s de tamaño n se puede tomar como una variable aleatoria n-dimensional $(X_1, X_2, \ldots X_n)$

Muestra	(X_1, X_2)	$P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1) \cdot P(X_2 = x_2)$
(B,B)	(0,0)	$\frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$
(B,A)	(0,1)	$\frac{1}{3} \cdot \frac{1}{6} = \frac{1}{18}$
(B,R)	(0,2)	$\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$
(A,B)	(1,0)	$\frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18}$
(A,A)	(1,1)	$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$
(A,R)	(1,2)	$\frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$
(R,B)	(2,0)	$\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$
(R,A)	(2,1)	$\frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$
(R,R)	(2,2)	$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Si interpretamos la m.a.s. como variable aleatoria bidimensional, la distribución de probabilidad sería:

$X_1 \backslash X_2$	0	1	2
0	1/9	1/18	1/6
1	1/18	1/36	1/12
2	1/6	1/12	1/4

Ejemplo.
Supongamos que la variable aleatoria X de una población se distribuye según una distribución de Poisson de parámetro λ, $X \rightsquigarrow P(\lambda)$, siendo:

$$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}.$$

Si tomamos una m.a.s. de tamaño n, hallar la función de distribución de la muestra.
Si tomamos una muestra de tamaño n, $X_1, X_2, \ldots X_n$ donde $X_i \rightsquigarrow P(\lambda)$

Entonces:

$$P(X_1=x_1, X_2=x_2, \dots, X_n=x_n) = \prod_{i=1}^{n} P(X=x_i) = \prod_{i=1}^{n} \left(\frac{e^{-\lambda}\cdot \lambda^{x_i}}{x_i!}\right)$$

$$= \frac{e^{-\lambda}\cdot\lambda^{x_1}}{x_1!}\cdot\frac{e^{-\lambda}\cdot\lambda^{x_2}}{x_2!}\cdots\cdots\frac{e^{-\lambda}\cdot\lambda^{x_n}}{x_n!} = \frac{e^{-\lambda\cdot n}\cdot\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$ esta

es la función de probabilidad de una muestra aleatoria de una población de Poisson.

Ejemplo

Supongamos que la variable aleatoria X se distribuye según una distribución exponencial de parámetro λ. (es una distribución continua). La función de densidad de esta distribución es:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x>0 \\ 0 & \text{si } x\leq 0 \end{cases}$$

La media es $\mu = E[X] = \frac{1}{\lambda}$

La varianza es $Var[X] = \frac{1}{\lambda^2}$

Si extraemos una muestra aleatoria simple de esta población $(x_1, x_2, \dots x_n)$, entonces la función de densidad de esta muestra es:

$$f(X_1=x_1, X_2=x_2, \dots, X_n=x_n) = f(X_1=x_1 \cap X_2=x_2 \cap \dots \cap X_n=x_n)$$

Al ser las variables independientes se tiene que:

$$f(X_1=x_1 \cap X_2=x_2 \cap \dots \cap X_n=x_n) = f(X_1=x_1)\cdot f(X_2=x_2)\cdots f(X_n=x_n)$$

Luego:

$$f(X_1=x_1, X_2=x_2, \dots, X_n=x_n) = \prod_{i=1}^{n} f(X_i=x_i) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} =$$

$$= \lambda\cdot e^{-\lambda x_1}\cdot\lambda\cdot e^{-\lambda x_2}\cdots\cdots\lambda\cdot e^{-\lambda x_n} = \lambda^n\cdot e^{-\lambda\sum x_i}$$

Ejemplo

Supongamos que la variable aleatoria X se distribuye según una binomial $X \rightsquigarrow B(k, p)$. Sabemos que la función de probabilidad viene dada por la expresión:

$$P[X=x] = \binom{K}{x} \cdot p^x \cdot q^{K-x} \quad \text{siendo } q = 1-p$$

Si tomamos una muestra aleatoria simple $(x_1, x_2, \ldots x_n)$ entonces la función de probabilidad de esta muestra es:

$$P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n] = \prod_{i=1}^{n} P[X_i = x_i) =$$

$$\prod_{i=1}^{n} \binom{K}{x_i} \cdot p^{x_i} \cdot q^{K-x_i} = \binom{K}{x_1} p^{x_1} \cdot q^{K-x_1} \cdot \binom{K}{x_2} \cdot p^{x_2} \cdot q^{K-x_2} \ldots$$

$$\ldots \binom{K}{x_n} \cdot p^{x_n} \cdot q^{K-x_n} = \left(\prod_{i=1}^{n} \binom{K}{x_i} \right) \cdot p^{\sum x_i} \cdot q^{\sum K-x_i} =$$

$$= \prod_{i=1}^{n} \binom{K}{x_i} \cdot p^{\sum x_i} \cdot q^{nK-\sum x_i}$$

Parámetros

Un parámetro es un valor numérico que describe una característica de una distribución. Se representan mediante letras griegas ($\lambda, \mu, \sigma, \theta$). Suelen ser valores desconocidos, son datos poblacionales.

Estadísticos

Llamaremos estadístico a cualquier función real de la variable muestral $(x_1, x_2, \ldots x_n)$ que no contiene ningún parámetro desconocido. Representaremos normalmente a los estadísticos con las letras T ó θ.

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$(x_1, x_2, \ldots x_n) \longrightarrow T(x_1, x_2, \ldots x_n)$$

Hemos de observar que a su vez T es una nueva variable aleatoria unidimensional, cuyo valor depende de la muestra $(x_1, x_2, \ldots, x_n)$ elegida. Tiene entonces pleno sentido preguntarse por la distribución que sigue T, sus momentos, su función generatriz de momentos, su media, su varianza, etc.

Veamos un ejemplo: sea X una v.a. poblacional que se distribuye con función de distribución $F(X, \theta)$ θ es un parámetro desconocido.
Tomamos muestras de tamaño n $(x_1, x_2, \ldots x_n)$ y definimos la función

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$(x_1, x_2 \cdots x_n) \longrightarrow T(x_1, x_2, \cdots x_n) = T = \frac{\sum_1^n x_i}{n}$$

Nos preguntamos, ¿es T un estadístico?
Sí por dos razones:
* T se construye a partir de una muestra
* ¿T tiene algún parámetro desconocido? No porque n es el tamaño de la muestra que queda conocido en cuanto definimos la muestra.

Luego T es un estadístico
Si hubiéramos definido T de la siguiente manera:

$T(x_1, x_2, \cdots x_n) = T = \dfrac{\sum_1^n x_i}{n} - n\theta$ no sería estadístico ya que contiene un parámetro desconocido θ.

Ejemplo
Sea X una variable aleatoria que sigue una distribución de Bernouilli de parámetro $p = 0'8$, es decir $X \rightsquigarrow b(0'8)$. Tratamos de obtener la distribución del estadístico media muestral, para muestras de tamaño $n = 2$.
La función de probabilidad de X es:

x	0	1
$P(X=x)$	0'2	0'8

Si tomamos muestras de tamaño 2 la función de probabilidad es:

(x_1, x_2)	$P(X_1 = x_1, X_2 = x_2)$
$(0,0)$	$0'2 \cdot 0'2 = 0'04$
$(0,1)$	$0'2 \cdot 0'8 = 0'16$
$(1,0)$	$0'8 \cdot 0'2 = 0'16$
$(1,1)$	$0'8 \cdot 0'8 = 0'64$

La variable aleatoria media muestral $\bar{X} = \frac{X_1 + X_2}{2}$ puede tomar los valores $0, \frac{1}{2}$ y 1. La probabilidad de cada uno de estos valores es:

$$P(\bar{X} = 0) = P(X_1 = 0, X_2 = 0) = 0'04$$

$$P(\bar{X} = \tfrac{1}{2}) = P(X_1 = 0, X_2 = 1 \cup X_1 = 1, X_2 = 0) =$$

$$= P(X_1 = 0, X_2 = 1) + P(X_1 = 1, X_2 = 0) - \cancel{P(X_1 = 0, X_2 = 1 \cap X_1 = 1, X_2 = 0)}^{=0}$$

$$= 0'16 + 0'16 = 0'32$$

$$P(\bar{X} = 1) = P(X_1 = 1, X_2 = 1) = 0'64$$

Por tanto la función de probabilidad de $\bar{X}$ es:

$\bar{X}$	0	$1/2$	1
$P(\bar{X} = x)$	$0'04$	$0'16 \cdot 2 = 0'32$	$0'64$

Estadísticos más usuales

1- Momentos muestrales respecto al origen (ordinarios)

Sea X una variable aleatoria de una población. Se define el momento de orden k con respecto al origen de X como:

$$\alpha_k = E[X^k], \quad k = 1, 2, \ldots,$$ este es un valor poblacional

Si tomamos muestras de tamaño n $(X_1, X_2, \ldots X_n)$, el estadístico muestral momento de orden k respecto al origen se define como:

$$O_k = \frac{\sum_{i=1}^{n} X_i^k}{n} \qquad k = 1, 2, 3, \ldots$$

El estadístico momento de orden k respecto al origen más importante es para $k = 1$, se trata de la media muestral:

$$O_1 = \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Para $k = 2$ obtenemos $O_2 = \frac{\sum_{i=1}^{n} x_i^2}{n}$

Momentos centrales

Sea X una v. a. en una población con media $\mu = E[X]$. Se define el momento central de orden k de la población como:

$$\mu_k = E[(X-\mu)^k] \quad k = 1, 2, \ldots$$

Si tomamos muestras de tamaño n en la población se define el estadístico muestral "momento central de orden k" como:

$$c_k = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^k}{n} \quad k = 1, 2, \ldots \text{ siendo}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Propiedades:

- Si $k = 1$ $\quad c_1 = 0$ ya que:

$$c_1 = \frac{\sum (x_i - \bar{x})}{n} = \frac{\sum x_i}{n} - \frac{\sum \bar{x}}{n} = \bar{x} - \frac{n\bar{x}}{n} = 0$$

- Si $k = 2 \Rightarrow c_2 = S^2$ (varianza muestral)

$c_2 = \frac{\sum (x_i - \bar{x})^2}{n} = S^2$. A efectos de cálculos también podemos poner que $c_2 = O_2 - O_1^2$ ya que

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{\sum x_i^2 - 2x_i\bar{x} + \bar{x}^2}{n} =$$

$$S^2 = \frac{\sum x_i^2}{n} - 2\bar{x} \cdot \frac{\sum x_i}{n} + \bar{x}^2 =$$

$$S^2 = \frac{\sum x_i^2}{n} - 2\bar{x}^2 + \bar{x}^2$$

$$S^2 = \frac{\sum x_i^2}{n} - \bar{x}^2 = O_2 - O_1^2$$

Cuasi-varianza muestral

La cuasi-varianza es un valor muestral, no tiene sentido el valor poblacional. Sea $x_1, x_2, \ldots x_n$ una m.a.s. de una población. Se define la cuasi varianza de la muestra como:

$$S_c = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Proposición

Sea $x_1, x_2, \ldots x_n$ una m.a.s. de una población. Entonces se verifica que:

$$n \cdot S^2 = (n-1) \cdot S_c^2$$

Demostración

$$\left.\begin{aligned} S^2 &= \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} \Rightarrow n \cdot S^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \\ S_c^2 &= \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \Rightarrow (n-1) S_c^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \end{aligned}\right\} \Rightarrow n S^2 = (n-1) S_c^2$$

Propiedades de los estadísticos más usuales

1º La esperanza que tiene el estadístico media muestral es la media de la población.

Demostración:

Tomemos una m.a.s. $x_1, x_2, \ldots x_n$. Como las x_i están identicamente distribuidas, como elementos de la

población verifican que $E[x_i] = \mu \quad \forall i = 1 \ldots n$.
La media muestral es:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Si tomamos la media:

$$E[\bar{x}] = E\left[\frac{\sum_{i=1}^{n} x_i}{n}\right] = \frac{1}{n} \cdot E\left[\sum_{i=1}^{n} x_i\right] =$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} E[x_i] = \frac{1}{n} \cdot \sum_{i=1}^{n} \mu = \frac{1}{n} \cdot n\mu = \mu$$

2ª La varianza de la media muestral es igual a la varianza de la población dividida entre el tamaño de la muestra

Demostración:

Tomemos una m.a.s $x_1, x_2, \ldots x_n$. Al estar las x_i identicamente distribuidas $\Rightarrow Var[x_i] = \sigma^2$ (varianza de la población). La media muestral es:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \Rightarrow Var[\bar{x}] = Var\left[\frac{\sum_{i=1}^{n} x_i}{n}\right] =$$

$= \frac{1}{n^2} \sum_{i=1}^{n} Var[x_i]$, al ser los x_i independientes la varianza de la suma es igual a la suma de las varianzas. Como $Var[x_i] = \sigma^2$ queda:

$$Var[\bar{x}] = \frac{1}{n^2} \cdot \sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

Teorema central del límite (versión general)

Sean $x_1, x_2, \ldots x_n$ variables aleatorias independientes e identicamente distribuidas. Si n es grande ($n \to \infty$) entonces:

$$\bar{x} \rightsquigarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \qquad (n \geq 30 \text{ son muestras grandes})$$

Media de la varianza muestral

Sea $X_1, X_2, \ldots, X_n$ una m.a.s. en una población. Sea S^2 la varianza muestral y sea σ^2 la varianza poblacional. Entonces:

$$E[S^2] = \frac{n-1}{n}\sigma^2$$

Demostración

La varianza muestral es

$$S^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

Podemos poner:

$$S^2 = \frac{\sum_{i=1}^{n}[(x_i-\mu)+(\mu-\bar{x})]^2}{n}$$ y desarrollando:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2+2(x_i-\mu)(\mu-\bar{x})+(\mu-\bar{x})^2}{n}$$, separamos

en tres sumas y si tenemos en cuenta que $2(\mu-\bar{x})$ es constante frente a i, $i=1\cdots n$, como también lo es $(\mu-\bar{x})^2=(\bar{x}-\mu)^2$:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} + \frac{2(\mu-\bar{x})}{n}\sum_{i=1}^{n}(x_i-\mu) + \frac{\cancel{n}\cdot(\bar{x}-\mu)^2}{\cancel{n}}$$

Como $\sum x_i-\mu = \sum x_i - \mu n = n\bar{x}-\mu n = -n(\mu-\bar{x})$, nos queda que:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} - 2\frac{(\mu-\bar{x})}{\cancel{n}}\cdot\cancel{n}\cdot(\mu-\bar{x})+(\bar{x}-\mu)^2$$

$$S^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} - 2(\bar{x}-\mu)^2+(\bar{x}-\mu)^2$$ es decir:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} - (\bar{x}-\mu)^2$$. Si tomamos esperanza en esta expresión obtenemos:

$$E[S^2] = \frac{\sum_{i=1}^{n} E[(x_i-\mu)^2]}{n} - E(\bar{x}-\mu)^2$$

Como: $E[(x_i-\mu)^2] = \sigma^2$, $E[\bar{x}-\mu] = \frac{\sigma^2}{n}$

$$E[S^2] = \frac{n \cdot \sigma^2}{n} - \frac{\sigma^2}{n} = \sigma^2 - \frac{\sigma^2}{n} = \frac{n\sigma^2 - \sigma^2}{n}$$

Por tanto

$$E[S^2] = \frac{(n-1)\sigma^2}{n}$$

Proposición: Esperanza de la varianza muestral

Sea $x_1, x_2, \dots x_n$ una m.a.s. en una población. Sea S^2 la varianza muestral y sea σ^2 la varianza poblacional. Entonces:

$$Var[S^2] = \frac{(n-1)^2}{n^3} \left(\mu^4 + \frac{n-3}{n-1} \cdot \sigma^4 \right) \text{ para } n > 1$$

Proposición: Esperanza de la cuasivarianza

Sea $x_1, x_2, \dots, x_n$ una m.a.s. en una población. Sea S_c la cuasi-varianza muestral

$$S_c = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Entonces $E[S_c] = \sigma^2$

Demostración:

Anteriormente hemos demostrado que $nS^2 = (n-1)S_c^2$
$\Rightarrow S_c^2 = \frac{n}{n-1} S^2$. Si tomamos esperanza:

$E[S_c^2] = \frac{n}{n-1} \cdot E[S^2]$. Según hemos probado
$E[S^2] = \frac{n-1}{n}\sigma^2$ y por tanto:

$$E[S_c^2] = \frac{n}{n-1} \cdot \frac{n-1}{n} \sigma^2 = \sigma^2$$

Distribución de la media muestral

a) Función de densidad de la media muestral

Proposición: Sea X una variable aleatoria con función generatriz de momentos $\phi_X(t)$. Sea $X_1, X_2, \ldots$ $\ldots X_n$ una m.a.s. Entonces la media muestral $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ tiene por función generatriz de momentos:

$$\phi_{\bar{X}}(t) = \left(\phi_X\left(\frac{t}{n}\right)\right)^n$$

Demostración: (No es necesario saberla)

Hay que tener en cuenta que al ser $X_1, X_2, \ldots X_n$ m.a.s. $\Rightarrow$ $X_1, X_2, \ldots X_n$ son independientes y verifican la propiedad:

$$E\left[\prod_{i=1}^{n} e^{\beta X_i}\right] = \prod_{i=1}^{n} E\left[e^{\beta X_i}\right]$$

Tenemos que $\phi_{\bar{X}}(t) = E\left[e^{t\bar{X}}\right] = E\left[e^{t\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)}\right]$

$$= E\left[e^{\frac{tX_1}{n}} \cdot e^{t\frac{X_2}{n}} \cdots e^{\frac{tX_n}{n}}\right] = E\left[\prod_{i=1}^{n} e^{t\frac{X_i}{n}}\right] = \prod_{i=1}^{n} E\left[e^{\frac{t}{n}X_i}\right]$$

Como $X_1, X_2, \ldots X_n$ son idénticamente distribuidas tienen la misma distribución de probabilidad y por tanto:

$$E\left[e^{\frac{t}{n}X_i}\right] = E\left[\frac{t}{n}X\right] \quad \forall i = 1 \ldots n. \text{ Por tanto:}$$

$$\phi_{\bar{X}}(t) = \prod_{i=1}^{n} E\left[e^{\frac{t}{n}X}\right] = \prod_{i=1}^{n} \phi_X\left(\frac{t}{n}\right) = \left(\phi_X\left(\frac{t}{n}\right)\right)^n$$

Ejemplo:

La función generatriz de momentos de una distribución gamma $\Gamma(p, a)$ es $\phi_X(t) = \left[1 - \frac{t}{a}\right]^{-p}$.

Sea $X_1, X_2, \ldots X_n$ un m.a.s. de una población que se distribuye según una $\Gamma(p, a)$. Nos piden que hallemos la función generatriz de momentos de la media muestral $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$

Si aplicamos la proposición anterior

$\phi_{\bar{X}}(t) = \left(\phi_X\left(\frac{t}{n}\right)\right)^n$. Como $\phi_X(t) = \left[1 - \frac{t}{a}\right]^{-p} \Rightarrow$

$\Rightarrow \phi_{\bar{X}}(t) = \left[1 - \frac{t}{na}\right]^{-np}$

b) Función de probabilidad/densidad de la variable aleatoria media muestral

Puede ser más o menos fácil determinar la función de probabilidad/densidad de la media muestral dependiendo de la situación. Para algunas variables cuya función de probabilidad/densidad conocemos, podemos determinar la función de probabilidad/densidad de la media muestral a partir del estadístico suma total. A veces se hace tan complicado obtener la función de probabilidad/densidad de la media muestral que no hay más remedio que echar mano del Teorema Fundamental del límite: si el tamaño de la m.a.s. $X_1, X_2, \dots X_n$ es grande ($n \geq 30$) entonces $\bar{X} \rightsquigarrow \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ donde μ es la media poblacional de X_i (están idénticamente distribuidas) y σ es la desviación típica de las X_i. Veamos algunos ejemplos.

Ejemplo 1

Supongamos que la variable X se distribuye según una distribución de Bernouilli de parámetro p, es decir $X \rightsquigarrow b(p)$. Sea $X_1, X_2, \dots X_n$ una m.a.s. Sabemos que $T = X_1 + X_2 + \dots X_n$, $T \rightsquigarrow B(n, p)$. Por tanto la función de probabilidad es

$P(T=k) = \binom{n}{k} p^k \cdot q^{n-k} \qquad q = 1 - p.$

La media muestral $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{T}{n}$ verifica:

$$P(\bar{X}=x) = P\left(\frac{T}{n}=x\right) = P(T=nx) = \binom{n}{nx} \cdot p^{nx} \cdot q^{n-nx}$$

es decir:

$$P(\bar{X}=x) = \binom{n}{nx} \cdot p^{nx} \cdot q^{n(1-x)}$$

Como $nx = 0, 1, 2, \dots n \Rightarrow x = 0, \frac{1}{n}, \frac{2}{n}, \dots 1$

Observación: La distribución binomial es reproductiva, es decir que si:

$$\left.\begin{array}{l} X_1 \rightsquigarrow B(n_1, p) \\ X_2 \rightsquigarrow B(n_2, p) \end{array}\right\} \Rightarrow X_1 + X_2 \rightsquigarrow B(n_1+n_2, p)$$

Ejemplo 2:

Sea X una variable aleatoria distribuida de esta manera $X \rightsquigarrow B(k, p)$. Tomamos una m.a.s $X_1, X_2, \dots, X_n$. Entonces la media muestral $\bar{X}$ $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ tiene la siguiente función de probabilidad.

$$P(\bar{X}=x) = \binom{nk}{nx} p^{nx} \cdot q^{n(k-x)}, \quad x = 0, \frac{1}{n}, \frac{2}{n}, \dots k$$

Vamos a demostrar esto:

Como $X_i \rightsquigarrow B(k, p)$ y la binomial es reproductiva $T = \sum_{i=1}^{n} X_i \rightsquigarrow B(nk, p)$. Por tanto la función de probabilidad de T es $P(T=t) = \binom{nk}{t} p^t \cdot q^{nk-t}$ donde $t = 0, 1, 2, \dots nk$. La función de probabilidad de la media muestral es:

$$P(\bar{X}=x) = P\left(\frac{T}{n}=x\right) = P(T=nx) = \binom{nk}{nx} p^{nx} \cdot q^{n(k-x)}$$

$nx = 0, 1, 2, \dots nk \Rightarrow x = 0, \frac{1}{n}, \frac{2}{n}, \dots k$

Observación: La distribución de Poisson de parámetro λ es reproductiva, es decir:

Si $\begin{matrix} X_1 \rightsquigarrow P(\lambda_1) \\ X_2 \rightsquigarrow P(\lambda_2) \end{matrix} \Rightarrow X_1 + X_2 \rightsquigarrow P(\lambda_1+\lambda_2)$

Ademas si $X \rightsquigarrow P(\lambda)$ la función de probabilidad (es de variable discreta) es:

$$P(X=x) = \frac{\lambda^x \cdot e^{-x}}{x!} \qquad x = 0, 1, 2, \dots$$

Ejemplo 3

Sea X una variable aleatoria distribuida segun una distribución de Poisson de parámetro λ. Tomemos una m.a.s $X_1, X_2, \dots X_n$. Entonces la función de probabilidad de la media muestral $\bar{X} = \frac{X_1+X_2+\dots+X_n}{n}$ es:

$$P(\bar{X}=x) = \frac{e^{-\lambda n} \cdot (n\lambda)^{nx}}{(nx)!}, \quad x = 0, \frac{1}{n}, \frac{2}{n}, \dots$$

Demostración:

Si $X_i \rightsquigarrow P(\lambda) \quad \forall i = 1 \dots n$, por la propiedad reproductiva de la distribución de Poisson:

$T = \sum_{i=1}^{n} X_i \rightsquigarrow P(n\lambda)$. Por tanto la función de probabilidad de T es:

$P(T=t) = \frac{e^{-n\lambda} \cdot (n\lambda)^t}{t!}$. Como $\bar{X} = \frac{T}{n}$, la función de probabilidad de $\bar{X}$ es:

$$P[\bar{X}=x] = P\left[\frac{T}{n}=x\right] = P[T=nx] = \frac{e^{-\lambda n} \cdot (n\lambda)^{nx}}{(nx)!}$$

$nx = 0, 1, 2, \dots \Rightarrow x = 0, \frac{1}{n}, \frac{2}{n}, \dots$

Observación: cambio de variable en variables aleatorias continuas

En los ejemplos anteriores hemos hecho cambios de variable, que en variable discreta son muy sencillos. En variable continua esto se complica un poco, hay que tener en cuenta lo siguiente:

Sea $f_X(x)$ la función de densidad de una variable continua X. Hacemos el cambio de variable $Y = g(X)$. Entonces para la función de densidad $f_Y(y)$ de Y se verifica que:

$$f_Y(y) = f_X(y) \cdot \left| \frac{dY}{dX} \right|$$

. Veamos un ejemplo:

Sea $X \rightsquigarrow Exp(\lambda)$. Entonces la función de densidad de X es:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en el resto} \end{cases}$$

Si hacemos el cambio de variable $Y = 4X - 1$ entonces la función de densidad de Y es:

$$f_Y(y) = f_X(y) \cdot \left| \frac{dY}{dX} \right| = 4 \cdot f_X(y)$$

como $y = 4x - 1 \Rightarrow x = \frac{y+1}{4}$. Por tanto:

$$f_Y(y) = 4 f_X(y) = \begin{cases} 4\lambda \cdot e^{-\lambda \cdot \frac{y+1}{4}} & \text{si } \frac{y+1}{4} > 0 \\ 0 & \text{en el resto} \end{cases}$$

es decir que:

$$f_Y(y) = \begin{cases} 4 \cdot \lambda \cdot e^{-\lambda \cdot \frac{y+1}{4}} & \text{si } y > -1 \\ 0 & \text{en el resto} \end{cases}$$

Observación: la variable aleatoria gamma

La variable aleatoria gamma de parámetro p y λ, $G(p, \lambda)$ tiene como función de densidad:

$$f(x) = \begin{cases} \dfrac{\lambda^p}{\Gamma(p)} x^{p-1} \cdot e^{-\lambda x} & \text{si } 0 < x < \infty,\ p > 0 \\ 0 \quad \text{si } x \leq 0 \end{cases}$$

$G(p,\lambda)$ es reproductiva, es decir que verifica:
Si $X_i \rightsquigarrow G(p_i, \lambda)$ $i = 1, 2, \ldots n$ entonces:

$$\sum_{i=1}^{n} X_i \rightsquigarrow G\left(\sum_{i=1}^{n} p_i, \lambda\right)$$

Cuando $p = 1$ obtenemos la distribución exponencial $Exp(\lambda) = G(1, \lambda)$ cuya función de densidad, si tenemos en cuenta que $\forall p \in \mathbb{Z}$ $\Gamma(p) = (p-1)! \Rightarrow \Gamma(1) = 0! = 1$, es:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en el resto} \end{cases}$$

$Exp(\lambda)$ no es reproductiva, pero como sí lo es $G(p, \lambda)$ podemos afirmar que si tomamos una m.a.s. de variable exponencial $X_1, X_2, \ldots, X_n$ entonces:

$$X_i \rightsquigarrow Exp(\lambda_i) \quad \forall i = 1, 2, \ldots n \Rightarrow \sum_{i=1}^{n} X_i \rightsquigarrow G(n, \lambda)$$

Volvamos a nuestros ejemplos.

Ejemplo 4

Sea X una v.a. tal que $X \rightsquigarrow Exp(\lambda)$. La función de densidad de X es entonces:

$$f_X(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en el resto} \end{cases}$$

Tomemos una m.a.s. $X_1, X_2, \ldots, X_n$. Según hemos visto si $T = \sum_{i=1}^{n} X_i \Rightarrow T \rightsquigarrow G(n, \lambda)$. Por lo tanto la función de densidad de T es:

$$f_T(t) = \begin{cases} \dfrac{\lambda^n}{(n-1)!}\, t^{n-1} \cdot e^{-\lambda t} & \text{si } 0 < t < \infty \\ 0 \text{ en el resto} \end{cases}$$

Lo que pretendemos es obtener la función de densidad de $\bar{X} = \frac{T}{n} \Rightarrow T = n\bar{x}$, este es el cambio de variable que necesitamos hacer. Según hemos visto en la observación sobre el cambio de variable, la función de densidad de $\bar{X}$ es:

$$f_{\bar{x}}(x) = f_T(x) \cdot \left| \frac{dX}{dT} \right| = \frac{1}{n} \cdot f_T(x) \cdot$$ Como $t = nx$,

cambiamos este valor en la función de densidad de T $f_T(t)$ y obtenemos la función de densidad de $\bar{x}$:

$$f_{\bar{x}}(x) = \frac{1}{n} \cdot f_T(t) = \begin{cases} \dfrac{1}{n} \cdot \dfrac{\lambda^n}{(n-1)!} \cdot (nx)^{n-1} \cdot e^{-\lambda(nx)} & \text{si } 0 < x < \infty \\ 0 \text{ en el resto} \end{cases}$$

es decir que la función de densidad de $\bar{x}$ es:

$$f_{\bar{x}}(x) = \begin{cases} \dfrac{\lambda^n}{n!}\, (nx)^{n-1} \cdot e^{-\lambda(nx)} & \text{si } x > 0 \\ 0 \text{ en el resto.} \end{cases}$$

Sabemos resolver la siguiente situación: tenemos una variable aleatoria X con distribución conocida. Tomamos una m.a.s. $X_1, X_2, \ldots X_n$ de forma que también conocemos la distribución de $T = \sum_{i=1}^{n} X_i$. Entonces mediante un cambio de variable llegamos a obtener la función de densidad de $\bar{X} = \sum_{i=1}^{n} X_i$.

¿Qué ocurriría si no conocemos la distribución de T? Esto es lo que vamos a ver a continuación. Supongamos que tomamos muestras de tamaño 2:

Sea X una v. a. con función de densidad $f(x)$. Tomemos una m.a.s. X_1, X_2 y supongamos que no tenemos ni puñetera idea de la distribución de $T = X_1 + X_2$. Entonces procedemos de la siguiente manera:

1º Seleccionamos la variable aleatoria bidimensional (W, U) donde $W = \dfrac{X_1 + X_2}{2}$ $\quad U = X_1$ ó bien $U = X_2$ (como queramos).

2º Hallamos la función de densidad de la variable (W, U), a la que llamaremos $g(W, U)$

3º Hallamos la función de densidad de W como función de densidad marginal de $g(W, U)$

A continuación vamos a ver algunos resultados que vamos a necesitar para dar los tres pasos anteriores

Observación: función de densidad de una variable aleatoria bidimensional.

Sea la variable aleatoria bidimensional (X_1, X_2) y supongamos que X_1 y X_2 son independientes. Sean $f_1(X_1)$ y $f_2(X_2)$ las funciones de densidad de X_1 y X_2 respectivamente. Entonces si $f(X_1, X_2)$ es la función de densidad conjunta de X_1 y X_2 se verifica que:

$$f(X_1, X_2) = f_1(X_1) \cdot f_2(X_2)$$

Observación: cambio de variable en variable aleatoria bidimensional.

Sea $f: \mathcal{D} \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ la función de densidad de una variable bidimensional $(X_1, X_2) \in \mathcal{D}$.

Sean las funciones cambio de variable:

$$h_1 : \mathcal{D}^* \longrightarrow \mathcal{D}$$
$$(W, U) \longrightarrow h(W, U) = X_1$$

$$h_2 : \mathcal{D}^* \longrightarrow \mathcal{D}$$
$$(W, U) \longrightarrow h_2(W, U) = X_2$$

Entonces la función de densidad de la variable bidimensional (W, U) es:

$$g(w,u) = f(h_1(w,u),\ h_2(w,u)) \cdot |J(x_1,x_2)|$$

donde $|J(x_1,x_2)|$ es el determinante de la matriz jacobiana de (x_1, x_2) ~~en~~ valor absoluto, es decir:

$$|J(x_1,x_2)| = \left| \begin{vmatrix} \frac{\partial x_1}{\partial w} & \frac{\partial x_1}{\partial u} \\ \frac{\partial x_2}{\partial w} & \frac{\partial x_2}{\partial u} \end{vmatrix} \right| = \left| \begin{vmatrix} \frac{\partial h_1}{\partial w} & \frac{\partial h_1}{\partial u} \\ \frac{\partial h_2}{\partial w} & \frac{\partial h_2}{\partial u} \end{vmatrix} \right|$$

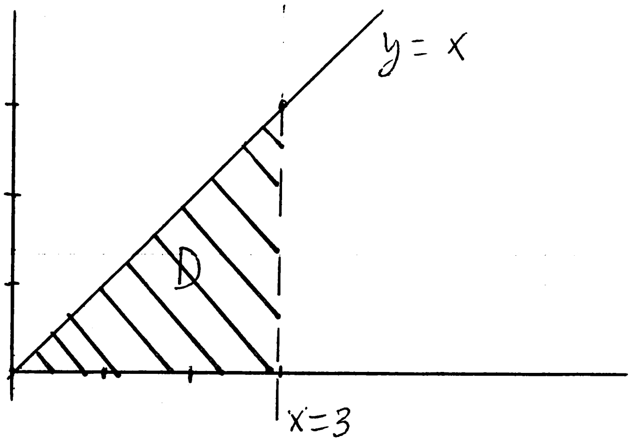
BREVE RECORDATORIO DE INTEGRALES DOBLES

Ejemplo:

Integrar $\iint_D (4x+y)\,dx\,dy$ siendo D el recinto

$$D = \{(x,y) \in \mathbb{R}^2 \,/\, 0 \le x \le 3,\ y \le x\}$$

En primer lugar dibujamos el recinto



Entonces:

$$\iint_D (4x+y)\,dx\,dy = \int_0^3 \left(\int_0^x (4x+y)\,dy \right) dx = \frac{81}{2}$$

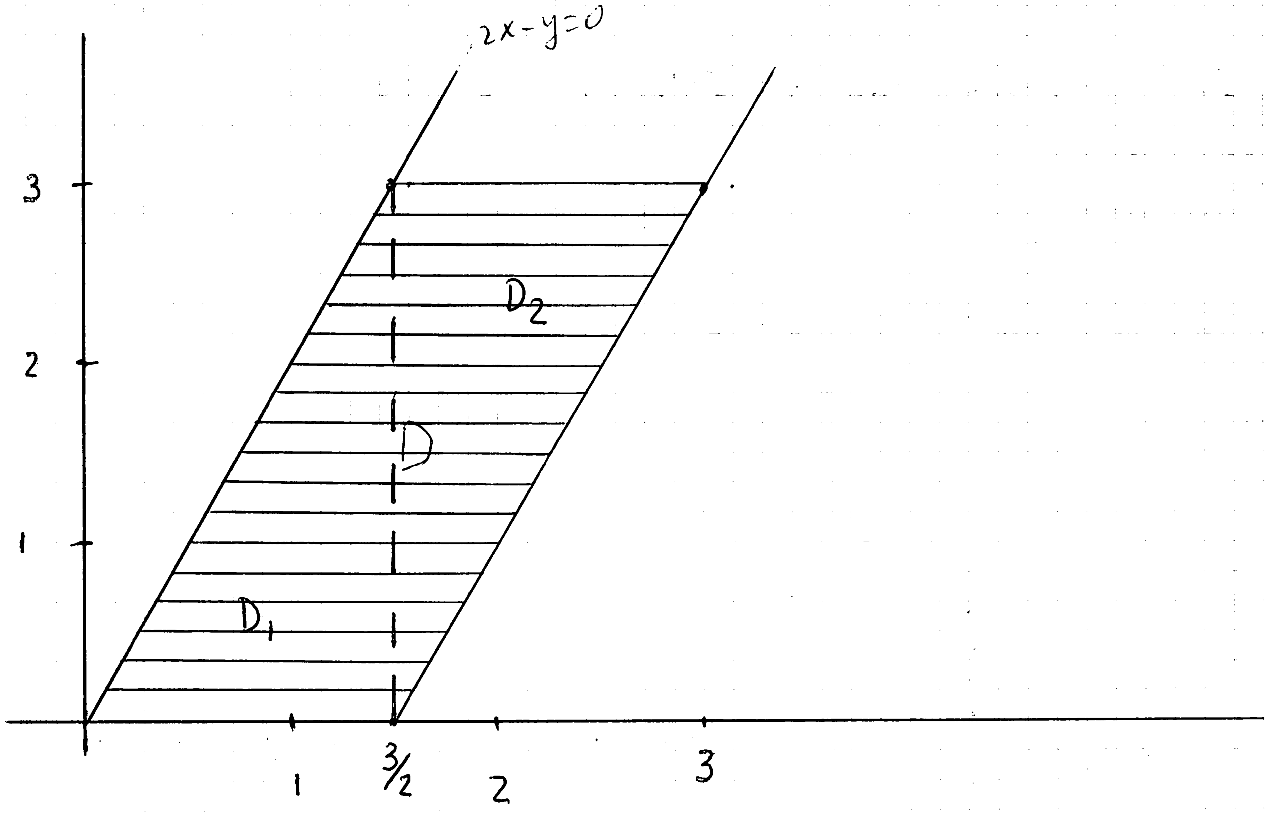
Tambien podriamos haber hecho variar primero y, y despues x:

$$\iint_D (4x+y)\,dx\,dy = \int_0^3 \left(\int_y^3 (4x+y)\,dx \right) dy = \frac{81}{2}$$

El recinto D es un recinto simple. Los recintos compuestos son aquellos en los que una de las variables, o las dos, pueden tomar más de dos valores. Para integrar dominios compuestos se dividen estos en recintos simples. Veamos un ejemplo:

$$\iint_D 7\,dx\,dy \quad \text{siendo} \quad D = \{(x,y) \in \mathbb{R}^2 \;/\; 0 \le y \le 3,\ 0 \le 2x - y \le 3\}$$

En primer lugar dibujamos el recinto.



El recinto D no es simple ya que:

* Si $0 \le x \le \frac{3}{2} \Rightarrow 0 \le y \le 2x \longrightarrow$ Recinto 1 : D1

* Si $\frac{3}{2} \le x \le 3 \Rightarrow 2x-3 \le y \le 3 \longrightarrow$ Recinto 2 : D2

Dividimos el recinto D en los recintos D_1 y D_2. Resolvemos la integral del siguiente modo:

$$\iint_D 7\,dx\,dy = \int_0^{3/2} \left(\int_0^{2x} 7\,dy \right) dx + \int_{3/2}^{3} \left(\int_{2x-3}^{3} 7\,dy \right) dx$$

Ejemplo 5:

Sea X una variable aleatoria continua tal que $X \rightsquigarrow U(0,3)$. Tomemos una m.a.s. de tamaño 2 x_1, x_2. El problema es determinar la función de densidad de la media muestral $\frac{x_1+x_2}{2}$. La variable continua $X \rightsquigarrow U(0,3)$ no es reproductiva. No sabemos que distribución sigue x_1+x_2. La viariable $X \rightsquigarrow U(0,3)$ tiene por función de densidad:

$$f(x) = \begin{cases} \frac{1}{3-0} & \text{si } 0 \leq x \leq 3 \\ 0 & \text{en el resto.} \end{cases}$$

Como $x_1 \rightsquigarrow U(0,3)$ $x_2 \rightsquigarrow U(0,3)$ las funciones de densidad de x_1 y x_2 son:

$$f_1(x_1) = \begin{cases} \frac{1}{3} & \text{si } 0 \leq x_1 \leq 3 \\ 0 & \text{en el resto} \end{cases} \qquad f_2(x_2) = \begin{cases} \frac{1}{3} & \text{si } 0 \leq x_2 \leq 3 \\ 0 & \text{en el resto.} \end{cases}$$

Al ser x_1 y x_2 independientes, si $f(x_1, x_2)$ es la función de densidad de (x_1, x_2) entonces:

$$f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2) = \begin{cases} \frac{1}{9} & \text{si } 0 \leq x_1 \leq 3,\ 0 \leq x_2 \leq 3 \\ 0 & \text{en el resto} \end{cases}$$

Sea $W = \frac{x_1+x_2}{2}$ $\quad u = x_1 \Rightarrow 2W = x_1 + x_2 \Rightarrow$

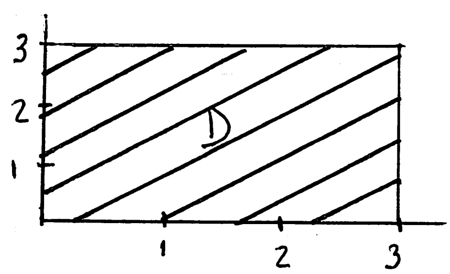
$\Rightarrow 2W = u + x_2 \Rightarrow 2W - u = x_2$

El cambio de variable que vamos a hacer es

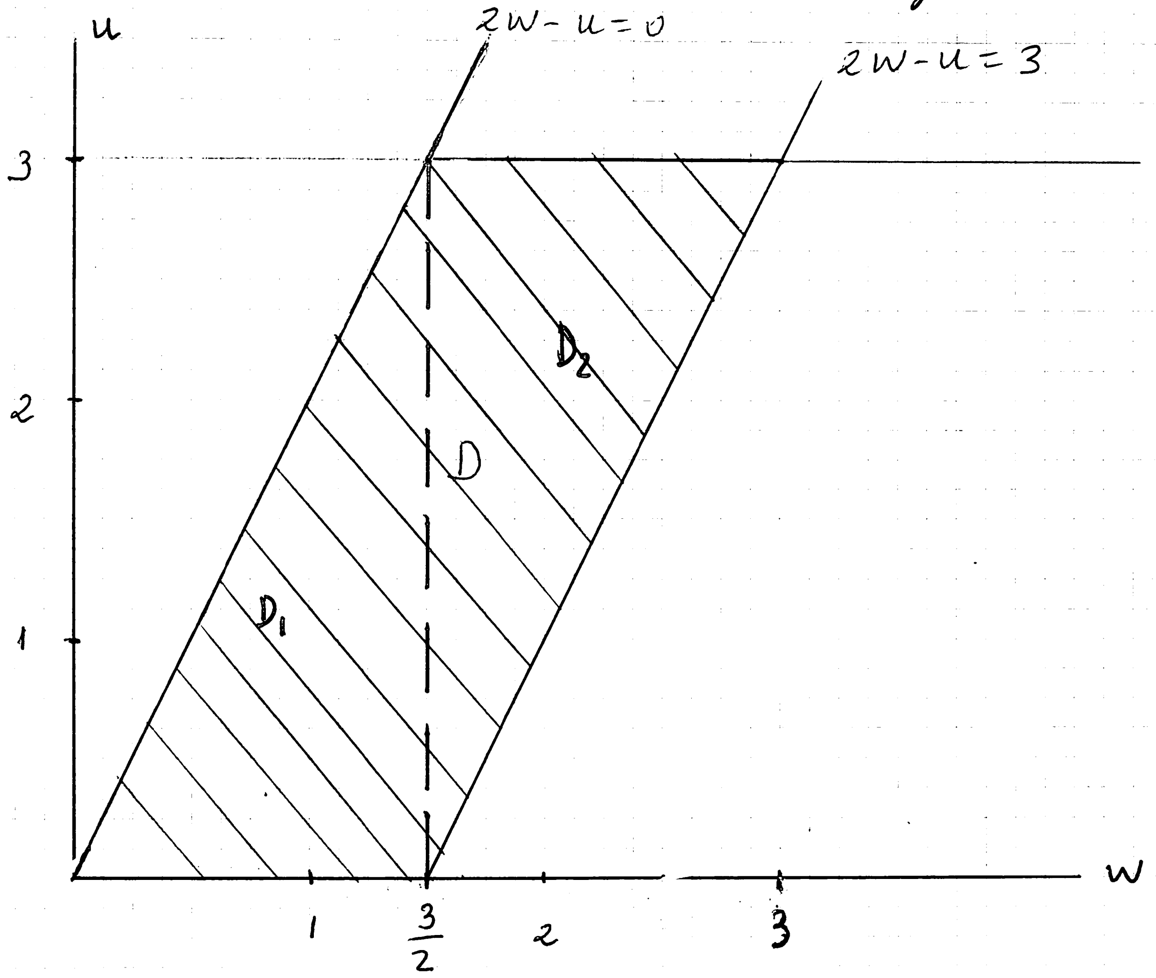
$x_1 = u \Rightarrow h_1(W,u) = u = x_1$

$x_2 = 2W - u \Rightarrow h_2(W,u) = 2W - u = x_2$

El dominio $\mathcal{D}$ de (x_1, x_2) es: $\mathcal{D} = (0,3) \times (0,3)$



El dominio $\mathcal{D}^*$ de (w,u) vamos a dibujarlo



La función de densidad $g(w,u)$ de la variable (w,u) es:

$$g(w,u) = f(h_1(w,u), h_2(w,u)) \cdot \left| \, |J(x_1,x_2)| \, \right|$$

$$|J(w,u)| = \begin{vmatrix} \frac{\partial h_1}{\partial w} & \frac{\partial h_2}{\partial u} \\ \frac{\partial h_2}{\partial w} & \frac{\partial h_2}{\partial u} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 2 & -1 \end{vmatrix} = -2$$

Entonces:

$$g(w,u) = 2 \cdot \begin{cases} \frac{1}{9} & \text{si } 0 < u < 3 \, , \; 0 < 2w - u < 3 \\ 0 & \text{en el resto} \end{cases}$$

Luego

$$g(w,u) = \begin{cases} \frac{2}{9} & \text{si } 0 < u < 3 \, , \; 0 < 2w - u < 3 \\ 0 & \text{en el resto} \end{cases}$$

Ahora sólo tenemos que hallar la densidad marginal de w, $g_1(w)$. Se tiene que:

$$g_1(w) = \int_{-\infty}^{\infty} g(w,u)\,du = \begin{cases} \int_0^{2w} \frac{2}{9}\,du & \text{si } 0 < w < \frac{3}{2} \\ \int_{2w-3}^{3} \frac{2}{9}\,du & \text{si } \frac{3}{2} < w < 3 \\ 0 & \text{en el resto} \end{cases}$$

es decir que la función de densidad de la media muestral $w = \overline{X} = \frac{X_1+X_2}{2}$ es:

$$g_1(w) = \begin{cases} \frac{4}{9}w & \text{si } 0 < w < \frac{3}{2} \\ (6-2w)\cdot\frac{2}{9} & \text{si } \frac{3}{2} < w < 3 \\ 0 & \text{en el resto} \end{cases}$$

Es muy importante conocer la función de densidad $f(x)$ de la variable media muestral $\overline{X}$, pues con ella podemos calcular:

* Función de distribución $F(x) = \int_{-\infty}^{x} f(t)\,dt$

* Media $E[X] = \int_{-\infty}^{+\infty} x\,f(x)\,dx$, etc, etc.

El método del ejemplo anterior también sirve para hallar la función de densidad de cualquier otro estadístico.

Ejemplo 6:

Sea $X \rightsquigarrow U(0,3)$. Tomemos una m.a.s. de dimensión 2 X_1, X_2. Se trata de hallar la función de densidad de $Z = X_1 \cdot X_2$. Seguimos el mismo procedimiento que en el ejemplo 5:

1º Hallamos la función de densidad $f(x_1, x_2)$ de la variable bidimensional (X_1, X_2) aprovechando que X_1 y X_2 son independientes.

$f(x_1, x_2) = f_1(x_1)\cdot f_2(x_2)$

Como $X_1 \rightsquigarrow U(0,3)$ y $X_2 \rightsquigarrow U(0,3)$ al igual que en el caso anterior

$$f(x_1,x_2) = f_1(x_1)\cdot f_2(x_2) = \begin{cases} \frac{1}{9} & \text{si } 0<x_1<3,\ 0<x_2<3 \\ 0 & \text{en el resto} \end{cases}$$

2º Sea $W = X_1 \cdot X_2$ $\quad U = X_1 \Rightarrow X_2 = \frac{W}{U}$ $\quad (u = x_1 \neq 0)$

El cambio de variable que hacemos es:

$x_1 = u = h_1(w,u)$

$x_2 = \frac{w}{u} = h_2(w,u)$

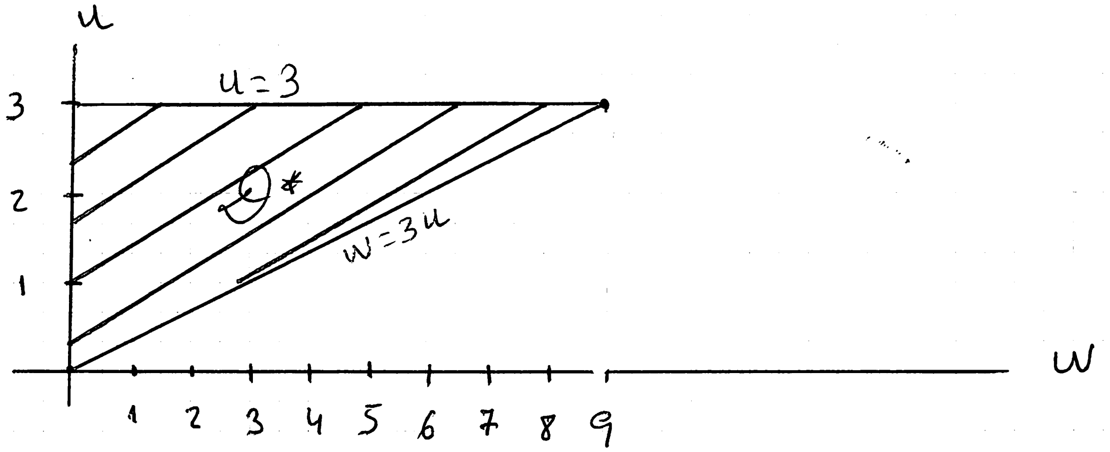
El determinante jacobiano del cambio es:

$$J(x_1,x_2) = \begin{vmatrix} \frac{\partial x_1}{\partial w} & \frac{\partial x_1}{\partial u} \\ \frac{\partial x_2}{\partial w} & \frac{\partial x_2}{\partial u} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \frac{1}{u} & -\frac{w}{u^2} \end{vmatrix} = -\frac{1}{u}$$

Veamos cual es el dominio $\mathcal{D}^*$ de (w,u), que dibujaremos

$\mathcal{D}^* = \{(w,u) \in \mathbb{R}^2 \ / \ 0<u<3,\ 0<\frac{w}{u}<3\}$ es decir

$\mathcal{D}^* = \{(w,u) \in \mathbb{R}^2 / \ 0<u<3,\ 0<w<3u\}$ al ser $u>0$.



La función de densidad $g(w,u)$ de la variable bidimensional (w,u) es:

$$g(w,u) = f[h_1(w,u), h_2(w,u)] \cdot \left| -\frac{1}{u} \right|$$

$$g(w,u) = \begin{cases} \frac{1}{9u} & \text{si } 0 < u < 3, \quad 0 < w < 3u \\ 0 & \text{en el resto} \end{cases}$$

Lo que nos interesa hallar ahora es la función de densidad marginal de w, $g_1(w)$

Se tiene que:

$$g_1(w) = \int_{-\infty}^{+\infty} g(w,u)\, du = \begin{cases} \int_{\frac{w}{3}}^{3} \frac{1}{9u}\, du & \text{si } 0 < w < 9 \\ 0 & \text{en el resto} \end{cases}$$

$$g_1(w) = \begin{cases} \frac{1}{9} \ln\left(\frac{9}{w}\right)^{(*)} & \text{si } 0 < w < 9 \\ 0 & \text{en el resto} \end{cases}$$ esta es la

función de densidad de $w = x_1 \cdot x_2$

(*)

$$\int \frac{1}{9u}\, du = \frac{1}{9} \ln u$$

$$\int_{\frac{w}{3}}^{3} \frac{1}{9u}\, du = \left[\frac{1}{9} \ln u \right]_{\frac{w}{3}}^{3} = \frac{1}{9} \ln 3 - \frac{1}{9} \ln \frac{w}{3} =$$

$$= \frac{1}{9} \ln \frac{3}{\frac{w}{3}} = \frac{1}{9} \ln \frac{9}{w}$$

Para acabar todo el repertorio de posibilidades que se nos pueden presentar para calcular las funciones de probabilidad/densidad de una variable aleatoria, volvamos a la variable discreta para ver como se resuelve el siguiente ejemplo.

Ejemplo 7:

Sea X una v.a. cuya función de probabilidad es:

x	0	1	2
$P[X=x]$	0'5	0'2	0'3

Pretendemos hallar la función de probabilidad de la media muestral. El problema es bien distinto a los anteriores, a parte de que X no sigue una función de probabilidad estandarizada, tampoco conocemos la función de probabilidad de la suma total.

Sea X_1, X_2 una m.a.s. de tamaño 2. Las funciones de probabilidad de X_1 y X_2 (por estar identicamente distribuidas) son:

x_1	0	1	2
$P[X_1=x_1]$	0'5	0'2	0'3

x_2	0	1	2
$P[X_2=x_2]$	0'5	0'2	0'3

Por ser X_1 y X_2 independientes, la función de probabilidad de la variable (X_1, X_2) se calcula como:

$$P[X_1=x_1, X_2=x_2] = P[X_1=x_1] \cdot P[X_2=x_2]$$

(X_1, X_2)	$P[X_1=x_1, X_2=x_2)$	$\bar{X} = \frac{X_1+X_2}{2}$
(0,0)	0'25	0
(0,1)	0'1	1/2
(0,2)	0'15	1
(1,0)	0'1	1/2
(1,1)	0'04	1
(1,2)	0'06	3/2
(2,0)	0'15	1
(2,1)	0'06	3/2
(2,2)	0'09	2

Ahora hallamos la función de probabilidad de $\bar{X}$

x	0	$\frac{1}{2}$	1	$\frac{3}{2}$	2
$P[\bar{X}=x]$	$0'25$	$0'2$	$0'34$	$0'12$	$0'09$

$P[\bar{X}=\frac{1}{2}] = P(0,1) + P(1,0) = 0'1 + 0'1 = 0'2$

$P[\bar{X}=1] = P(0,2) + P(1,1) + P(2,0) = 0'15 + 0'04 + 0'15 = 0'34$

$P[\bar{X}=\frac{3}{2}] = P(1,2) + P(2,1) = 0'6 + 0'6 = 0'12$

$P[\bar{X}=2] = P(2,2) = 0'09$

Insistimos de nuevo, una vez que conocemos la función de probabilidad, podemos calcular la función de distribución de la población, la esperanza, los momentos, etc.

Muestras de población normal

Lema de Fisher - Cochram

Sea X una variable con distribución $X \rightsquigarrow N(\mu, \sigma)$
Sea $X_1, X_2, \dots X_n$ una m.a.s. Sean $\bar{X}$ y S^2 la media y la varianza muestral. Entonces:

i) $\bar{X}$ y S^2 son independientes (lo cual es equivalente a que también lo sean $\bar{X}$ y S_c^2)

ii) $\bar{X} \rightsquigarrow N(\mu, \frac{\sigma}{\sqrt{n}})$

iii) $\frac{nS^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1}$ (o lo que es equivalente

$$\frac{(n-1)\,S_c^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1}$$

Ejemplo:
Sea $X \rightsquigarrow N(\mu = 5, \sigma^2 = 0'01)$. Consideremos una m.a.s. de tamaño 16. Hallar:

a) $P(5 < \bar{x} < 5'2)$

b) $P(S^2 < 0'019125)$

Solución:

a) Según el lema de Fisher-Cochran $\bar{x} \rightsquigarrow N(\mu, \frac{\sigma}{\sqrt{n}})$ es decir que:

$$\bar{x} \rightsquigarrow N\left(5, \frac{\sqrt{0'01}}{\sqrt{16}}\right) \Rightarrow x \rightsquigarrow N\left(5, \frac{0'1}{4}\right)$$

$$\bar{x} \rightsquigarrow N\left(5, \frac{1}{40}\right) \Rightarrow \frac{(x-5)}{\frac{1}{40}} \rightsquigarrow N(0,1)$$

$$P(5 < \bar{x} < 5'2) = P\left(\frac{5-5}{\frac{1}{40}} < \frac{\bar{x}-5}{\frac{1}{40}} < \frac{5'2-5}{\frac{1'}{40}}\right)$$

$$= P(0 < Z < 8) = P(Z < 8) - P(Z < 0) = 1 - \frac{1}{2} = \frac{1}{2}$$

b) De nuevo, según el teorema de Fisher-Cochran

$$\frac{nS^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1} \Rightarrow \frac{16 \cdot S^2}{0'01} \rightsquigarrow \chi^2_{15} \Rightarrow$$

$$\Rightarrow 1.600\, S^2 \rightsquigarrow \chi^2_{15}$$

Vamos a recordar brevemente la definición de las distribuciones Ji-cuadrado de Pearson, F de Snedekor y T de Student.

Distribución Ji-cuadrado de Pearson

Se dice que la variable aleatoria Z se distribuye según una Ji-cuadrado con n-grados de libertad y se escribe $Z \rightsquigarrow \chi^2_n$ si $Z = Z_1 + Z_2 + \cdots + Z_n$, donde $Z_i = X_i^2$ $\forall i = 1 \dots n$, $X_i \rightsquigarrow N(0,1)$ x_i independientes.

Esquemáticamente podemos poner $\chi^2_n = \sum_{i=1}^{n} [N(0,1)]^2$

La función χ^2_n es reproductiva, es decir:

Si $x_1 \rightsquigarrow \chi^2_{n_1}$, $x_2 \rightsquigarrow \chi^2_{n_2}$ $\Rightarrow$ $x_1 + x_2 \rightsquigarrow \chi^2_{n_1+n_2}$

Distribución F de Snedekor

Se dice que la variable aleatoria Z se distribuye según una F de Snedekor de m filas y n columnas o con (m, n) grados de libertad y se escribe $Z \rightsquigarrow F_{m,n}$ si:

$$Z = \frac{\dfrac{X_1}{m}}{\dfrac{X_2}{n}} \quad \text{siendo} \quad \left.\begin{array}{l} X_1 \rightsquigarrow \chi^2_m \\ X_2 \rightsquigarrow \chi^2_n \end{array}\right\} \begin{array}{l} X_1 \text{ y } X_2 \text{ son} \\ \text{independientes} \end{array}$$

Esquemáticamente podemos poner $F_{m,n} = \dfrac{\dfrac{\chi^2_m}{m}}{\dfrac{\chi^2_n}{n}}$

La distribución F no es reproductiva.

Distribución T de Student

Se dice que la variable aleatoria Z se distribuye según una T de Student con n grados de libertad y se escribe $Z \rightsquigarrow t_n$ si:

$$Z = \frac{X_1}{\sqrt{\dfrac{X_2}{n}}} \quad \text{donde} \quad \left.\begin{array}{l} X_1 \rightsquigarrow N(0,1) \\ X_2 \rightsquigarrow \chi^2_n \end{array}\right\} \begin{array}{l} X_1 \text{ y } X_2 \text{ son} \\ \text{independientes} \end{array}$$

Esquemáticamente podemos poner:

$$t_n = \frac{N(0,1)}{\sqrt{\dfrac{\chi^2_n}{n}}}$$

A continuación vamos a ver como se distribuyen las medias muestrales y las varianzas muestrales dependiendo de que se conozcan o no los parametros de la población. Hemos definido las distribuciones anteriores para demostrar las afirmaciones que hagamos.

Distribución de la media muestral de una población con varianza conocida: Inferencia sobre la media poblacional μ

Sea $X \rightsquigarrow N(\mu, \sigma)$ con varianza poblacional conocida σ^2. Tomemos una m.a.s. $x_1, x_2, \dots, x_n$. Entonces la media muestral verifica $\bar{x} \rightsquigarrow N(\mu, \frac{\sigma}{\sqrt{n}})$

Dem:

Este resultado es una parte del lema de Fisher. Con ello podemos hacer inferencia sobre el valor de la media de la población μ (normalmente desconocida) ya que $\bar{x}$ sí es conocida.

Inferencia sobre la media poblacional μ siendo la varianza poblacional σ^2 desconocida.

Sea $X \rightsquigarrow N(\mu, \sigma)$ con varianza desconocida σ^2. Tomemos una m.a.s $x_1, x_2, \dots, x_n$. Entonces:

$$\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}} \rightsquigarrow t_{n-1}$$

Dem:

Según el lema de Fisher se verifica:

$$\bar{x} \rightsquigarrow N(\mu, \frac{\sigma}{\sqrt{n}}) \Rightarrow \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \rightsquigarrow N(0,1)$$

$$\frac{n s^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1} \Rightarrow \frac{n \cdot s^2}{\sigma^2} \Big/ (n-1) \rightsquigarrow \frac{\chi^2}{n-1}$$

Como $t_{n-1} = \dfrac{N(0,1)}{\sqrt{\dfrac{\chi^2_{n-1}}{n-1}}}$

$$\frac{\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{\frac{n s^2}{\sigma^2}}{n-1}}} \sim t_{n-1} \Rightarrow \frac{\bar{x}-\mu}{\frac{\frac{\sigma}{\sqrt{n}} \cdot \frac{\sqrt{n}}{\sigma} \cdot s}{\sqrt{n-1}}} \rightsquigarrow t_{n-1} \Rightarrow$$

$$\Rightarrow \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n-1}}} \rightsquigarrow t_{n-1}$$

Inferencia sobre σ siendo μ conocida

Sea la variable aleatoria $X \rightsquigarrow N(\mu, \sigma)$, con μ conocida. Tomemos una m.a.s. $x_1, x_2, \dots x_n$. Entonces $x_i \rightsquigarrow N(\mu, \sigma)$ $\forall i = 1, \dots, n$ con lo que podemos afirmar que $\frac{x_i-\mu}{\sigma} \rightsquigarrow N(0,1)$. Por tanto:

$$\sum_{i=1}^{n} \left(\frac{x_i-\mu}{\sigma}\right)^2 \rightsquigarrow \chi^2_n \Rightarrow \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^2} \rightsquigarrow \chi^2_n$$

es decir que:

$$\frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} (x_i-\mu)^2 \rightsquigarrow \chi^2_n$$

Inferencia sobre σ siendo μ desconocida

Sea X una variable aleatoria normalmente distribuida con media desconocida μ y varianza también desconocida σ^2

$X \rightsquigarrow N(\mu, \sigma)$.

Si tomamos una muestra aleatoria simple de tamaño n: $x_1, x_2, \dots x_n$, por el lema de Fisher se verifica que:

$$\frac{n\,S^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1}$$

siendo S^2 la varianza de la muestra es decir

$$S^2 = \frac{\sum_{i=1}^{n} (x_i-\bar{x})^2}{n}$$

Los grados de libertad en la distribución χ^2 indican más precisión

Inferencia sobre $\mu_x - \mu_y$ con varianzas conocidas

Sean las v.a. X e Y distribuidas del siguiente modo:

$X \rightsquigarrow N(\mu_x, \sigma_x)$ $\quad Y \rightsquigarrow N(\mu_y, \sigma_y)$ siendo X e Y independientes.

Tomamos una m.a.s. de tamaño n en X, $X_1, X_2, \dots X_n$

Tomamos una muestra de tamaño m en Y, $Y_1, Y_2, \dots Y_m$

Entonces:

$$\bar{X} - \bar{Y} \rightsquigarrow N\left(\mu_x - \mu_y, \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}\right)$$

Demostración:

Por el lema de Fisher sabemos que se verifica:

$$\bar{X} \rightsquigarrow N\left(\mu_x, \frac{\sigma_x}{\sqrt{n}}\right) \qquad \bar{Y} \rightsquigarrow N\left(\mu_y, \frac{\sigma_y}{\sqrt{n}}\right)$$

Como:

$E[\bar{X} - \bar{Y}] = E[\bar{X}] - E[\bar{Y}] = \mu_x - \mu_y$, al ser X e Y independientes se tiene que:

$$Var[\bar{X} - \bar{Y}] = Var[\bar{X}] + Var[\bar{Y}] = \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}$$

y por tanto la desviación típica de $\bar{X} - \bar{Y}$ es:

$$\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}$$. luego:

$$\bar{X} - \bar{Y} \rightsquigarrow N\left(\mu_x - \mu_y, \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}\right)$$