TEMA 1:

DISTRIBUCIONES DE PROBABILIDAD FUNDAMENTALES

1.1. DISTRIBUCIONES DISCRETAS.

1.1.1. Binomial (1, p) o de Bernoulli.

Es una distribución auxiliar que toma dos valores:

1: que se identifica con la aparición del suceso.

0: que se identifica con la aparición del suceso contrario.

Sea A un suceso tal que $P(A) = p$, sabiendo que $0 \leq p \leq 1$.

$$X : \begin{cases} 1 \; ; \textit{si ocurre } A \\ 0 \; ; \textit{si NO ocurre } A \end{cases}$$

$$\boxed{p + q = 1}$$

Función de cuantía: $P(\xi = x) = p^x \cdot q^{1-x}$

Esperanza: $E(\xi) = p$

Varianza: $V(\xi) = p \cdot q$

El máximo valor de la varianza es 0´25, momento de máxima incertidumbre pues la probabilidad de éxito es la mismo que la probabilidad de fracaso p = q.

Desviación típica: $\sigma = +\sqrt{p \cdot q}$

1.1.2. Binomial (n, p).

Se denomina así por ser el resultado de sumar "n" variables aleatorias IID $B(1,p)$, es decir, $\xi = \xi_1 + \xi_2 + \ldots\ldots + \xi_n$ donde $\xi_i \to B(1,p)$, luego pondremos $\xi = \sum_{i=1}^{n} \xi_i \to B(n,p)$

F. Cuantía:

$$P(\xi = x) = \binom{n}{x} p^x q^{n-x} = \binom{n}{x} p^x (1-p)^{n-x} \qquad \forall\, x = 0, 1, \ldots\ldots, n.$$

Esperanza: $E(\xi) = n \cdot p$

Varianza: $V(\xi) = n \cdot p \cdot q$

Desviación típica: $\sigma = +\sqrt{n \cdot p \cdot q}$

Reproductividad, Suma o Adición:

La suma de dos o más binomiales independientes con igual probabilidad de éxito (p), es igual a otra binomial con la misma probabilidad de éxito y cuyo parámetro "n" es la suma de las "n" de las binomiales que sumemos.

$\xi_1 \sim B(n_1, p)$ y $\xi_2 \sim B(n_2, p)$ $\quad$ *Si* ξ_1 *y* ξ_2 *son independientes* $\quad \eta = \xi_1 + \xi_2 \sim B(n_1 + n_2, p)$

Ejemplo 1:

$\xi_1 \sim B(10, 0'3)$ y $\xi_2 \sim B(5, 0'3)$; *Si* ξ_1 *y* ξ_2 *son independientes* $\quad \eta = \xi_1 + \xi_2 \sim B(15, 0'3)$

Ejemplo 2:

$\xi_1 \sim B(10, 0'3)$ y $\xi_2 \sim B(5, 0'4)$ No se pueden sumar al tener distinta probabilidad de éxito.

1.1.3. Poisson (λ).

Poisson determinó esta distribución como límite de la B(n, p)

Siempre que $\begin{cases} n \to \infty \\ p \to 0 \quad \{ n \cdot p = \lambda \end{cases}$ $\qquad \xi \sim P(\lambda)$

La aproximación es buena si $n > 100$ y $p \leq 0'1$

Nos va a representar el número de sucesos en un espacio de tiempo determinado. La variable tiene un único parámetroλ. $x \sim P(\lambda) \qquad \lambda > 0$

F. Cuantía: $\quad P(\xi = x) = e^{-\lambda} \dfrac{\lambda^x}{x!} \quad para \quad x = 0,1,\ldots\ldots,\infty$

Esperanza y Varianza: $\quad E(\xi) = V(\xi) = \lambda$

Desviación típica: $\quad \sigma = +\sqrt{\lambda}$

Reproductividad, Suma o Adición:

La suma de dos o más variables independientes de Poisson, es otra Poisson cuyo parámetro λ se obtiene mediante la suma de los parámetros λ de las Poisson que se sumen.

$\xi_1 \sim P(\lambda_1)$ y $\xi_2 \sim P(\lambda_2) \quad Si\ \xi_1\ y\ \xi_2\ son\ independientes \quad \eta = \xi_1 + \xi_2 \sim P(\lambda_1 + \lambda_2)$

1.2. DISTRIBUCIONES CONTINUAS.

1.2.1. Uniforme o Rectangular.

Se dice que una variable ξ sigue la distribución rectangular o uniforme cuando se encuentra definida en el intervalo (a, b), con función de densidad (dicha función se utilizará para el cálculo de probabilidades, pues este modelo no posee tablas):

$$f(x) = \frac{1}{b-a} \quad para \;\; a \le x \le b \quad \rightarrow (b > a)\; \xi \rightarrow U(a,b)$$

Esperanza: $E(\xi) = \dfrac{b+a}{2}$

Varianza: $V(\xi) = \dfrac{(b-a)^2}{12}$

Desviación típica: $\sigma = +\sqrt{\dfrac{(b-a)^2}{12}}$

La distribución uniforme no posee propiedad reproductiva.

1.2.2. Normal (0, 1).

Es sin duda, la distribución de probabilidad más importante utilizada en economía por dos razones:

- Muchas variables de tipo económico siguen un comportamiento Normal.
- En muestras grandes ($n > 100$) las características aditivas (como la media poblacional) siguen una distribución casi normal.

Además es la base para la construcción de distribuciones en el muestreo imprescindibles para la inferencia estadística, tales como Chi-cuadrado, t-Student, F-Snedecor.

Esta Normal posee tablas para calcular probabilidades. $X \equiv N(0,1)$

1.2.3. Normal General.

Esta normal no posee tablas, para calcular cualquier probabilidad, debemos de transformarla en una normal tipificada "tiene tablas".

A la transformación se le conoce como tipificación de la variable:

$\xi \sim N(\mu,\sigma)$ (Normal generalizada)

$$Z = \frac{\xi - \mu}{\sigma}$$

$Z \sim N(0,1)$ (Normal tipificada)

Anagrama: $\xi \sim N(\mu,\sigma)$

1.2.4. Distribución Chi-Cuadrado de Pearson. $\chi^2(n)$

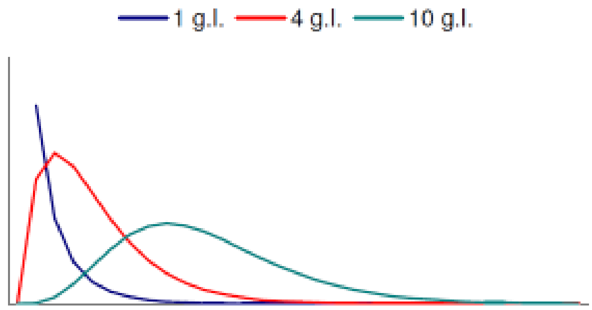
Esta distribución se utiliza para caracterizar la varianza muestral.

Se define como la suma de cuadrados de Normales reducidas independientes $\chi_n^2 = Z_1^2 + Z_2^2 + \ldots + Z_n^2$.
Recibe el nombre $\chi^2(n)$ siendo "n" el número de variables aleatorias que lo integran denominados grados de libertad.

Como el campo de variación de las variables normales es el intervalo $(-\infty, +\infty)$ y la variable χ_n^2 es suma de sus cuadrados, su campo de variación es $[0, +\infty)$.

La función de densidadde esta variable depende del número de variables que lo formen (número de grados de libertad), cuanto mayor sean los grados de libertad, menor será la altura de la campana.

Su forma depende de los grados de libertad.
Es asimétrica a la derecha y tiende a hacerse simétrica a medida que aumenta "n"
No tiene correspondencia alguna con la realidad.



<u>Esperanza:</u> $E(\chi_n^2) = n$

<u>Varianza:</u> $V(\chi_n^2) = 2n$

<u>Propiedad aditiva o reproductiva:</u>
Estas variables se pueden sumar, siempre que sean independientes

$$\left.\begin{array}{l} \chi_n^2 \\ \chi_m^2 \\ \chi_k^2 \end{array}\right\} \rightarrow \chi_{n+m+k}^2$$

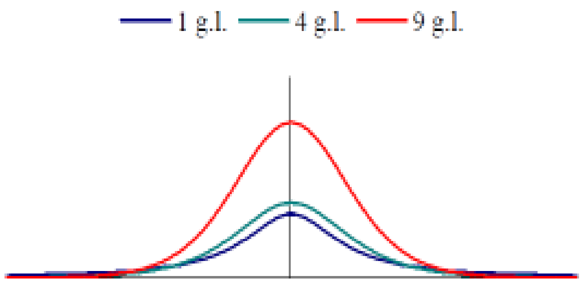
1.2.5. Distribuciónt-Student. $t(n)$

Esta distribución es el cociente entre una Normal reducida y la raíz cuadrada de una Chi cuadrado dividida por sus grados de libertad (siempre que las dos variables sean independientes), definimos la variable t-Student con n grados de libertad $t(n)=\dfrac{N(0,1)}{\sqrt{\dfrac{\chi_n^2}{n}}}$.

El número de grados de libertad es igual al número de variables que figuran en el denominador. El campo de variación de la variable es de $(-\infty,+\infty)$. Es simétrica (media = mediana = moda) respecto al eje de ordenadas y su forma depende de los grados de libertad, "n". Es muy parecida a la Normal pero no es exactamente la campana de Gauss. La t- Student tiene una mayor dispersión que la normal, por ello tiene un mayor número de valores centrales.
No tiene correspondencia alguna con la realidad.

<u>Esperanza:</u> $E(t_n)=0$

<u>Varianza:</u> $V(t_n)=\dfrac{n}{n-2}$; $(n>2)$



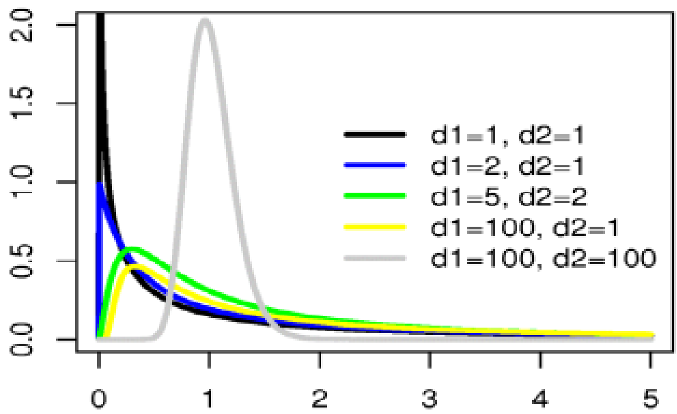
Para los valores elevadores de n (grados de libertad mayores que 30), esta distribución se aproxima a Normal

$$t_{n>30} \xrightarrow{n\to\infty} N(0,1)$$

1.2.6. Distribución F de Fisher-Snedecor. $F_{m,n}$

Consideramos $m+n$ variables aleatorias, $N(0,1)$ e independientes. $\eta_1, \eta_2, \eta_3, \ldots, \eta_m$; $\xi_1, \xi_2, \xi_3, \ldots, \xi_n$, entonces la variable aleatoria con m y n grados de libertad

$$F_{m,n} = \frac{\frac{\eta_1^2 + \eta_2^2 + \eta_3^2 + \ldots + \eta_m^2}{m}}{\frac{\xi_1^2 + \xi_2^2 + \xi_3^2 + \ldots + \xi_n^2}{n}} = \frac{\frac{\chi_m^2}{m}}{\frac{\chi_n^2}{n}}$$ recibe el nombre F de Fisher-Snedecor



No es simétrica.

Su campo de variación, como procede de una suma de cuadrados es de $[0, +\infty)$

Tiene una asíntota en $+\infty$.

No tiene correspondencia alguna con la realidad.

Una propiedad de F-Snedecor (propiedad de la variable), desde el punto de vista de las probabilidades es

$$\eta \to F(m,n) \quad \frac{1}{\eta} = F(n,m)$$

Una F-Snedecor con grados de libertad 1 y n es igual a t-student.

$$\mathrm{F}(1,\mathrm{n}) = t^2(n)$$

1.3. TEOREMA CENTRAL DEL LÍMITE.

La suma de un numero suficientemente grande ($n > 100$) de variables independientes se comporta aproximadamente como una normal.

Sean x_i "n" variables independientes, $Y=\sum_{i=1}^{n} x_i$ diremos que se aproxima (no que es, pues no se exige que la variables "x_i" sean normales) a una normal del tipo $Y \sim N\left(\sum_{i=1}^{n} E(x_i), \sqrt{\sum_{i=1}^{n} V(x_i)}\right)$

Ejercicio 1 Sean x_i para $i=1,2,\ldots\ldots,250$ variables independientes todas ellas con media 4 y desviación típica 3. Obtener la distribución de probabilidad de $Y=\sum_{i=1}^{250} x_i$

$$E(Y)=E\left(\sum_{i=1}^{250} x_i\right)=\sum_{i=1}^{250} E(x_i)=250\cdot 4=1.000$$

$$V(Y)=V\left(\sum_{i=1}^{250} x_i\right)=\sum_{i=1}^{250} V(x_i)=250\cdot 9=2.250$$

Luego diríamos $Y \sim N\left(\sum_{i=1}^{n} E(x_i), \sqrt{\sum_{i=1}^{n} V(x_i)}\right) = N\left(1.000, \sqrt{2.250}\right)$

Ejercicio 2. La demanda de un producto oscila diariamente entre 20 y 40 unidades determinar la probabilidad de que en un periodo de 182 días, el número de unidades demandadas supere las 6.370 unidades supuesto la independencia de la demanda de cada día respecto de los restantes.

$\xi_i \to$ Demanda diaria de un producto $\xi_i \to U(20,40)\begin{cases} E(\xi_i)=\dfrac{b+a}{2}=\dfrac{40+20}{2}=30 \\ V(\xi_i)=\dfrac{(b-a)^2}{12}=\dfrac{(40-20)^2}{12}=33'33 \end{cases}$

$$\eta=\sum_{i=1}^{182}\xi_i \to N\left(E(\eta),\sqrt{V(\eta)}\right)\begin{cases} E(\eta)=E\left(\sum_{i=1}^{182}\xi_i\right)=182\cdot 30=5.460 \\ V(\eta)=V\left(\sum_{i=1}^{182}\xi_i\right)=182\cdot 33'33=6.066'666 \end{cases}$$

$$\eta \to N\left(5.460,\sqrt{6.066'06}\right)$$

$$P(\eta>6370)\xrightarrow{\text{Tipificamos}}P\left(\frac{\eta-n\cdot\mu}{\sqrt{n\cdot\sigma^2}}>\frac{6370-5460}{\sqrt{6066'666}}\right)=P(Z>11'68)=0$$

1.- Una compañía de seguros tiene 1000 asegurados en el ramo de accidentes. Si la probabilidad de sufrir un accidente en un año para un asegurado cualquiera es de 0´005, el modelo mejor para el número de siniestros en un año es:

a) Normal (5; 2´23). b) Binomial (1000; 0´005).c) Chi-cuadrado. d) Poisson con λ=5.

2.-La demanda diaria de refrescos en una cafetería se distribuye uniformemente entre 1000 y 2000 unidades. Entonces:

a) Su valor medio es de 1500.
b) La probabilidad de que se demanden menos de 1750 unidades es de 0,25.
c) La probabilidad de que se demanden más de 1750 unidades es de 0,75.
d) Ninguna de las anteriores.

3.-Dada una distribución binomial de dos parámetros, B (10 ; 0´5), elegir la afirmación correcta.

a) La variable puede tomar cualquier valor menor que 10.
b) La media es 0,5.
c) Media y varianza coinciden.
d) Ninguna de las anteriores.

4.-Dada una distribución de Poisson, elegir la afirmación falsa.

a) Media y varianza coinciden.
b) Tiene un sólo parámetro.
c) La media sólo puede tomar valores enteros.
d) La variable nunca toma valores negativos.

5.- Elegir la afirmación correcta sobre la distribución normal.

a) Es una distribución discreta.
b) La media siempre será positiva.
c) Los valores de la variable aleatoria no pueden ser negativos.
d) Ninguna de las anteriores afirmaciones es verdad.

6.- ¿Cuál de las siguientes afirmaciones sobre el T.C.L. es falsa?

a) Hace referencia a la convergencia en distribución hacia el modelo normal.
b) Necesita para su aplicación práctica una suma numerosa de variables aleatoriasindependientes.
c) Permite, bajo ciertas condiciones, aproximar la distribución binomial a la normal.
d) Permite la convergencia hacia cualquier modelo de probabilidad.

7.- Una compañía de seguros tiene 1000 asegurados en el ramo de accidentes. Si la probabilidad de sufrir un accidente en un año para un asegurado cualquiera es de 0´005, el fenómeno aleatorio que modeliza la mejor aproximación para el número de siniestros en un año sigue una distribución:

a) Normal.

b) Binomial (1;p).

c) Chi-cuadrado.

d) Poisson.

8.- Sobre la demanda de un producto sólo se sabe que oscila, diariamente, entre 1000 y 2000 unidades. La probabilidad de que se demanden entre 1250 y 1750 unidades es igual a:

a) 0´75. b) 0´25. c) 0´5. d) Ninguna de las anteriores.

9.- Para la variable aleatoria que representa el número de caras obtenidas al lanzar 4 veces una moneda perfecta, encontramos que:

a) Su valor medio es 2 y su varianza es 0´5.

b) Sigue un modelo de Poisson.

c) Su media es cuatro.

d) Su varianza es 1.

10.- En una cadena de montaje se obtienen 10000 unidades de un artículo. Se conoce que la probabilidad de que una unidad sea defectuosa es de 0,001. Entonces, para la variable aleatoria que mide el número de unidades defectuosas encontramos que:

a) La probabilidad de que haya cinco defectuosas es de 0,7755.

b) Su media es el doble que su varianza.

c) Una buena aproximación es un modelo de Poisson.

d) Ninguna de las anteriores.

11.- El consumo diario de litros de café en un bar sigue una distribución N(100;25). La probabilidad de que en un día concreto se consuman exactamente 115,5 litros es igual a:

a) 0´6. b) 0´4.c) 0´25. d) Ninguna de las anteriores.

12.- Por investigaciones previas, se estima que la probabilidad de que una persona haga deporte más de 2 horas a la semana es de 0´15. En función de esto, la probabilidad de que en un grupo de 10 individuos haya 4 que hagan deporte es:

a) 0´0401. b) 0´1298.c) 0´0085. d) Ninguna de las anteriores.

SOLUCIONES: **1b, 2a, 3d, 4c, 5d, 6d, 7d, 8c, 9d, 10c, 11d, 12a,**