TEMA 65.- DISTRIBUCIONES DE PROBABILIDAD DE VARIABLE DISCRETA. CARACTERÍSTICAS Y TRATAMIENTO. DISTRIBUCIONES BINOMIAL Y POISSON. APLICACIONES

Índice

1. INTRODUCCIÓN

Supongamos un experimento que consiste en elegir al azar un individuo de la comunidad autónoma X.
Obviamente el espacio muestral $\Omega = \{w_i\}_{i=1}^{N}$ está formado por toda la población, es decir, por N individuos.
Dado que la elección es al azar, todos tienen la misma probabilidad de ser elegidos, i.e.

$$P(\{w_i\}) = \frac{1}{N}$$

Sin embargo, no solemos tener interés en la persona, en cambio sus características sí. Por ejemplo: su edad, peso, pensamiento político etc. En otras palabras, nuestro interés está en un valor numérico asociado a w_i, que denotamos

$$X(w)$$

Cabe señalar que ω es un resultado de un experimento aleatorio. Dado ω, entonces

$$X(\omega) = x \quad , \text{ viene dado.}$$

Ejemplos.-

a) Elegimos al azar una persona, entonces su edad ya no es aleatoria es un valor numérico fijo

b) Elegimos al azar una porción de residuos de una planta de reciclaje, entonces su composición viene dada.

Así, podemos deducir que nuestro objetivo es estudiar características numéricas de nuestro espacio de probabilidad, no dicho espacio en sí mismo, ya que por su abstracción es complicado.

2. VARIABLES ALEATORIAS

La única forma que conocemos de trasladar información de un espacio a otro es mediante una aplicación.

Luego, a la aplicación

$$X : \Omega \longrightarrow \mathbb{R}$$
$$\omega \longrightarrow X(\omega) = x$$

que asigna un valor numérico a los posibles resultados de un experimento se le llama variable aleatoria (v.a.)

Ejemplos.-

i) El nº de llamadas a una centralita en un periodo de tiempo

ii) Suma de las caras de dos dados al ser lanzados

Las razones fundamentales para llevar a cabo esta traslacion de los datos son:

1.) $\mathbb{R}$ es un espacio bien conocido. En comparación, con $(\Omega, \mathcal{A}, P)$ que es muy abstracto y poco cómodo para trabajar

2.) Fijar nuestra atención en el valor numérico nos permite construir un modelo probabilístico aplicable a Ω_S con rasgos similares

Así, hay que destacar que, aunque sean características numéricas concretas, una vez realizado el experimento, al estar asociadas a experimentos aleatorios se puede considerar a las v.a. como

"cantidades aleatorias"

<u>Def</u>^{n}.- Consideremos $(\Omega, \mathcal{A})$ un espacio medible y $(\mathbb{R}, \beta)$ otro con β la σ-álgebra de Borel. Una v.a. es una aplicación

$X : \Omega \longrightarrow \mathbb{R}$ que cumple:

$$X^{-1}(B) \in \mathcal{A}, \forall B \in \beta$$

<u>Lema</u>.- La familia de sucesos

$$\sigma(X) = \{X^{-1}(B) : B \in \beta\}$$

es una σ-álgebra que verifica que

$$\sigma(X) \subseteq \mathcal{A}$$

<u>Def</u>^{n}.- Dada una v.a. X, entonces se llama σ-álgebra inducida por X a

$$\sigma^{ind}(X) = \{X^{-1}(B) : B \in \beta\}$$

Además, puesto que tenemos una relación entre $(\Omega, \mathcal{A})$ y $(\mathbb{R}, \beta)$ por medio de la σ-álgebra inducida $\sigma(X)$, veamos que si $(\Omega, \mathcal{A}, P)$ es un espacio de probabilidad entonces X induce una probabilidad P_X sobre $(\mathbb{R}, \beta)$ de la forma siguiente

- $X^{-1}(B) \in \mathcal{A}, \quad \forall B \in \beta$
- Como $(\Omega, \mathcal{A}, P)$ es un espacio de probabilidad

$$P(X^{-1}(B)) =: P_X(B) \quad \forall B \in \beta$$

Entonces P_X hereda las propiedades de ser una función de probabilidad de P. Luego, $(\mathbb{R}, \beta, P_X)$ es un espacio de probabilidad.

Nota.- La probabilidad P_X también recibe el nombre de distribución de la v.a. X

Ejemplo.- Es importante señalar que P_X hereda las propiedades de P a través de X. Es decir, supongamos lanzamos 2 dados y definimos:

- X: suma de las caras $\Rightarrow P_X(\{0\}) = P(X^{-1}(0)) = P(\emptyset) = 0$
- Y: valor absoluto de la diferencia de las caras $\Rightarrow P_Y(\{0\}) = P(Y^{-1}(0)) = P(\{(1,1),(2,2),\dots(6,6)\}) = \frac{6}{36} = \frac{1}{6}$

Luego $P_X \neq P_Y$ por la def$^{\underline{n}}$ de X e Y

Sin embargo, de la misma forma P_X es un objeto matemático complejo y que requiere conocer

$$X^{-1}(B) \quad \forall B \in \beta$$

Algo que no ocurre, habitualmente, en la realidad. Así, recurrimos a diversas funciones que nos permiten describir y caracterizar la v.a. X

2.1. Función de distribución de una v.a

Defⁿ - Dada una v.a. X definimos su función de distribución como:

$$F_X(x) = P_X([-\infty, x]) = P(X^{-1}((-\infty, x])) =$$
$$= P(X \leq x) \quad \forall x \in \mathbb{R}$$

Propiedades - Dada un función de distribución de una v.a. Esta función cumple;

1. No negativa
2. Monótona creciente
3. Continua por la derecha
4. $\lim\limits_{x \to +\infty \ (-\infty)} F_X(x) = 1 \ (0)$

A dicha función se le llama también función de distribución acumulada puesto que;

$$P(a < X \leq b) = F(b) - F(a)$$

Observación.- Hemos definido F_X a partir de P_X. Sin embargo, el recíproco también es cierto. Así, dada la distribución de X o la distribución acumulada F_X podemos obtener la otra a través de la que tenemos.

2.2 V. a. Discreta

Una v. a. cuyo rango de valores posibles es finito o infinito numerable, se conoce como v. a. discreta

Ejemplo. - La edad de un individuo elegido al azar toma un rango de posibles valores infinito numerable, ya que es un entero positivo

$\Rightarrow$ La edad es una v. a. discreta.

- El número de la cara de un dado lanzado al azar toma $A := \{1, 2, 3, 4, 5, 6\}$ un valor

$x \in A$

Def^n.- Una v. a. discreta es aquella que verifica que $\exists D$ cjto numerable tal que

$$P(X \in D) = P_X(D) = 1.$$

- Es decir tenemos un espacio muestral numeble.

Así, si $x \in D^c \Rightarrow x \notin D \Rightarrow P_X(x) = 0$

$\Rightarrow \boxed{P(X = x) = 0}$

Def^n.- (Función de probabilidad o cuantía) La función de probabilidad o cuantía de una v. a. discreta, cuyo soporte es D, se define:

$$f_X(x) = \begin{cases} P(X = x), & \text{si } x \in D \\ 0, & \text{otro caso} \end{cases}$$

Proposición - Toda función de probabilidad cumple que:

i) f_X no negativa

ii) $\sum_{x_i \in D} f_X(x_i) = 1$

$\boxed{D}$ i) si $x \notin D$ soporte de $X \Rightarrow f_X(x) = 0$

si $x \in D \Rightarrow f_X(x) = P(X = x) \geq 0$

$\Rightarrow f_X$ no negativa

ii) $\sum_{x_i \in D} f_X(x_i) = \sum_{x_i \in D} P(X = x_i) \underset{\substack{\text{prop. prob} \\ \text{disj 2 a 2}}}{=} P(X \in D) = 1$

Nota - Notemos que

$$F_X(x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} f_X(x_i)$$

Entonces si ordenamos el soporte $D = \{x_{(i)}\}_{i \geq 1}$

Entonces

$$F_X(x_{(i)}) - F_X(x_{(i-1)}) = P(X = x_i) = f_X(x_i)$$

Luego es una función escalonada y cumple que si tenemos $F_X \Rightarrow f_X$ y viceversa

3. MODELOS DISCRETOS

En lo que sigue se estudian los ejemplos de v. a. discretas más importantes por su aplicación directa en la vida real

3.1. Distribución de Bernoulli

Habitualmente conocida como probabilidad éxito fracaso. Esto se debe a su característica binaria, luego el rango $\{0,1\}$ nos permite concluir que es v. a. discreta

Ejemplo - Escogemos un individuo al azar y nos preguntamos si es diabético, entonces

$$\{0 \text{ (no lo es)},\ 1 \text{ (lo es)}\}$$

Lanzamos una moneda, es cara o cruz ?

$\Rightarrow$ Binario $\Rightarrow$ v. a. Bernoulli

Como podemos definir su función de cuantía? En primer lugar, si $x \notin \{0,1\}$

$\Rightarrow P(X=x) = 0$. $\quad P(X=1) = p \in [0,1]$

$\quad P(X=0) = 1-p \in [0,1]$

Así,

$$f_X(x) = \begin{cases} p^x (1-p)^{1-x} & x \in \{0,1\} \\ 0 & x \notin \{0,1\} \end{cases}$$

es su función de cuantía

Si una v.a. discreta sigue esta distribución se dice que se distribuye como una Bernoulli y se denota por:

$$X \sim Be(p)$$

<u>Características</u>.- Si tomamos su valor esperado $E[X]$ y su $Var[X]$ varianza, veamos que

$$E[X] = \sum_{x \in \{0,1\}} x f_X(x) = f_X(1) = P(X=1) = p$$

$$Var[X] = E[(X-p)^2] = E[X^2 - 2Xp + p^2]$$
$$= E[X^2] - 2p\,E[X] + p^2$$
$$= E[X^2] - 2p^2 + p^2$$
$$= p - p^2 = p(1-p)$$

3.2. Distribución BINOMIAL

La distribución binomial se puede entender como una generalización de la dist. Benoulli. Supongamos $X_i \sim Be(p)$ con probabilidad de éxito p. Si repetimos n veces una prueba de Bernoulli de forma independiente una de otra, la v.a. que nos da el número de éxitos en las n pruebas, sigue una distribución binomial y se denota como

$$X \sim Bi(n, p),$$

Además

$$X = \sum_{i=1}^{n} X_i \; ; \quad X_i \in Be(p)$$

Notemos que la cantidad de éxitos varía entre $0, 1, \dots, n \Longrightarrow D = \{0, \dots, n\}$.

Def^n - Decimos que X es una v. binomial de parámetros n y p si su función de cuantía viene dada por

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x \in D \\ 0 & x \notin D = \{0, \dots, n\} \end{cases}$$

Observación → Supongamos $m \leq n$ éxitos

$\Longrightarrow \begin{cases} m \text{ éxitos} \rightsquigarrow p^m \\ (n-m) \text{ fracasos} \rightsquigarrow (1-p)^{n-m} \end{cases}$ y los m éxitos tienen n huecos para colocarlos $\rightsquigarrow \binom{n}{m}$

$$\Longrightarrow P(X=m) = \binom{n}{m} p^m (1-p)^{n-m}.$$

$\boxed{\text{Propiedad Fundamental de la Binomial}}$ Gracias a la expresión de $f_X(x)$ si $X \sim Bi(n,p)$

Entonces, se puede estimar el punto de inflexión de $f_X(x)$.

$\boxed{D}$

$$f_X(k+1) = \binom{n}{k+1} p^{k+1} (1-p)^{n-k-1}$$

$$= \frac{n! \, (n-k) \; p^k \cdot p \; (1-p)^{n-k-1} (1-p)}{(k+1)(n-k-1)! \cdot (n-k) \; (1-p)}$$

$$= \binom{n}{k} \frac{n-k}{k+1} p^k (1-p)^{n-k} \cdot \frac{p}{(1-p)}.$$

$$= f_X(k) \cdot \frac{n-k}{k+1} \frac{p}{1-p}$$

Entonces

$$f_X(k+1) > f_X(k) \iff \frac{n-k}{k+1}\,\frac{p}{1-p} > 1$$

Veamos si aislamos k, que obtenemos

$$\frac{n-k}{k+1} > \frac{1-p}{p} \iff n-k > (k+1)\,\frac{1-p}{p}$$

$$\iff n > k + k\left(\frac{1-p}{p}\right) + \frac{1-p}{p}$$

$$\iff p\left(n - \frac{1-p}{p}\right) > pk + k - kp$$

$$\iff np - 1 + p > k \iff \boxed{k < p(n+1) - 1}$$

Luego, $k \geq [p(n+1)-1] + 1 \rightsquigarrow f_X(k)$ decrece!

3.3. Distribución de Poisson

Esta distribución nos aparece ligada a sucesos o experimentos en los que queremos / interesa estudiar su ocurrencia en un periodo de tiempo finito, donde:

1. La probabilidad que el suceso ocurra es proporcional a la longitud del intervalo siendo λ la constante de proporcionalidad

$$p \simeq \lambda(b-a)$$

2. La ocurrencia + 2 de dos veces en un periodo corto de tiempo es casi nula.

Ejemplos.- El número de llamadas telefónicas recibidas por una centralita.

El número de accidentes en una pista de esquí.

La cantidad de estrellas fugaces observadas...

La distribución de Poisson de parámetro λ se denota como

$$X \sim P_0(\lambda)$$

El soporte D es

$$D = \{0, 1, \dots\}.$$

Además, la función de cuantía de X sigue la expresión:

$$f_X(x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!}, & x \in D \\ 0, & x \notin D \end{cases}$$

Así, $F_X(x) = \sum_{n \leq x} \dfrac{e^{-\lambda}\lambda^n}{n!}$

A continuación, enunciamos uno de los resultados que hacen a esta distribución tan importante

Prop.- Sea $\{X_n\}_{n\geq 1}$ una sucesión de v.a. tales que $X_i \sim Bi(n, p_n)$ cuyo $\lim_{n\to\infty} n p_n = \lambda$

Entonces

$$\lim_n P(X_n = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

La utilidad de dicho resultado reside en aproximar una $f_X(x)$ de una $X \sim Bi(n,p)$ por una

$$Y \sim P_0(\lambda), \quad \lambda = n \cdot p$$

Cuando n es muy grande y p muy pequeño

Por este motivo, se la conoce como la distribución de los sucesos raros.

Ejemplo.- Sea una comunidad autónoma cuyo $N = 350\,000$ y p la probabilidad de tener una enfermedad que afecta a 3.5 individuos por año $\Longrightarrow$

$p_c =$

$$p = \frac{1}{100.000} \quad y \quad n \cdot p = 3{,}5$$

Entonces, X = nº muertes por esta enfermedad

sigue $X \sim P_0(3{,}5)$

Características.- Calculemos $E[X]$ y $Var[X]$ de $X \sim P_0(\lambda)$

- $$E[X] = \sum_{x=0}^{+\infty} x\, f_X(x) = e^{-\lambda} \cdot \sum_{x=1} \lambda \frac{\lambda^{x-1}}{(x-1)!} =$$

$$= e^{-\lambda} \cdot \lambda \sum_{x=0} \frac{\lambda^x}{x!} =$$

$$= e^{-\lambda} e^{\lambda} \cdot \lambda = \lambda_{//}$$

- ~~$Var[X] = E[X(X-1)]$~~

- $$Var[X] = E[X^2] - (E[X])^2$$

$$= \underbrace{E[X^2] - E[X]} + E[X] - (E[X])^2$$

$$= E[X(X-1)] + \lambda - \lambda^2$$

$$E[X(X-1)] =$$

$$= e^{-\lambda} \sum_{x=1} x \cdot (x-1) \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=2}^{+\infty} \cancel{x} \cdot \cancel{(x-1)} \frac{\lambda^x}{\cancel{x!}}$$

$$= e^{-\lambda} \sum_{x=2}^{+\infty} \frac{\lambda^2 \lambda^{x-2}}{(x-2)!} = e^{-\lambda} e^{\lambda} \lambda^2 = \lambda^2$$

$$\Rightarrow Var[X] = \lambda^2 + \lambda - \lambda^2 = \lambda = E[X]$$

2h 13'

$\boxed{\text{4. CONCLUSIONES}}$

TEMA 66 DISTRIBUCIÓN DE PROBABILIDAD DE VARIABLE CONTINUA - CARACTERÍSTICAS Y TRATAMIENTO. DISTRIBUCIÓN NORMAL APLICACIONES

Índice

Empezamos en 2.2. Los apartados anteriores son idénticos al TEMA 65

2.2. V.A CONTINUAS

Una variable aleatoria continua (v.a.c), a diferencia de una v.a. discreta, es aquella que la probabilidad que la variable esté en un intervalo dado $[a,b]$ viene determinada por su integral sobre dicho intervalo

Nota.- Para desarrollar este tema es suficiente con asumir la siguiente definición.

Defⁿ.- Diremos que la v.a. es continua cuando existe una función no negativa, integrable y cuya integral sobre $\mathbb{R}$ es uno, que cumple:

$$\int_a^b f(x)\,dx = P(a < X \leq b) = P(X \in (a,b])$$

La función $f(x)$ anterior recibe el nombre de <u>función de densidad de probabilidad</u>

Además, dada la función de densidad, podemos obtener la función de distribución $F_X(x)$ de X v.a. continua por medio de:

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt$$

Luego si f es continua se tiene que

$$F'_X(x) = f(x)$$

<u>Nota</u>.- Observemos que $\forall X$ v.a. continua se satisface que

$$P(X=x) = \int_x^x f(t)\,dt = 0$$

Es decir, en v.a. continuas la función de densidad nos aporta cuanta densidad de probabilidad hay por segmento infinitesimal. Puesto que $X=x$ es un conjunto de medida $0 \Rightarrow P(X=x)=0$

Además,

$$P(X=x) + F_X(x-) = F_X(x)$$

$$\Rightarrow \quad F_X(x) = F_X(x-)$$

2.3. Esperanza de una v.a. continua

Supongamos que X v.a. continua no negativa con densidad f. Entonces, definimos su esperanza (también conocida como valor esperado o media) mediante la expresión siguiente:

$$E[X] = \int_{-\infty}^{+\infty} x\, f(x)\, dx = \int_{0}^{+\infty} x\, f(x)\, dx$$

En caso que NO sea no negativa se puede definir sobre X : $\begin{cases} x^+ = \max\{x, 0\} \\ x^- = \max\{-x, 0\} \end{cases}$

Entonces

$$EX = EX^+ - EX^-$$

Sin embargo, si la media existe habitualmente se usa la siguiente expresión:

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

Supongamos que transformamos X una $G(x)$ por medio de una función medible G. Entonces $G(X)$ es v.a.c por serlo X y entonces

$$E[G(x)] = \int_{-\infty}^{+\infty} g(x)\, f(x)\, dx$$

Propiedad - Sea X una v.a.c y g, h dos transformaciones de X, $a, b \in \mathbb{R}$. Entonces

$$E[a\,g(x) + b\,h(x)] = a\cdot E[g(x)] + b\,E[h(x)]$$

2.4. Momento central de una v.a.

A veces para conocer información acerca de X nos interesa conocer sus valores como de "alejados" estan del valor esperado.

Def^n - Dada una v.a. X definimos el momento central de orden K como

$$E\left[(X - EX)^K\right]$$

En particular, el momento central de orden $\boxed{K=2}$ se conoce como varianza

Sea X una v. continua entonces

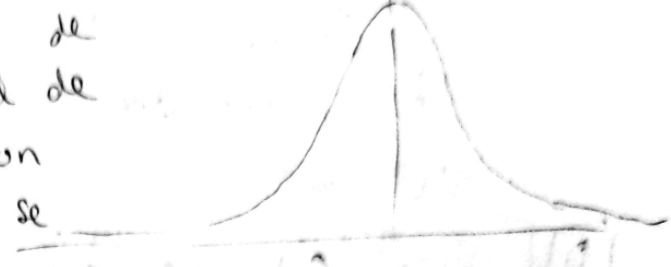
$$Var[x] = E\left[(X - EX)^2\right]$$

$$= \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,dx \quad , \quad \text{con } \mu = EX$$

3. DISTRIBUCIÓN NORMAL

La distribución normal es la piedra angular de la inferencia estadística al Análisis de Datos. Esto se debe a que las distribuciones de muchos "estadísticos muestrales" tienden a la distribución normal conforme aumenta el tamaño de la muestra.

La apariencia gráfica de dicha distribución es el de una curva simétrica con forma de campana que se extiende sin límite en las direcciones positivas y negativas.



La mencionada función de densidad fue descubierta por De Moivre en 1733. Posteriormente Laplace la estudió a fondo sin publicar algunos de sus resultados. Finalmente, en 1809, K. F. Gauss cuando estudiaba los errores cometidos en cualquier tipo de medición también la descubrió, publicando sus trabajos en dicho ámbito. Por ello, en ciertos contextos estadísticos se la conoce como Gaussiana.

Defn.- Se dice que una v.a X sigue una distribución normal (o simplemente, es normal) de parámetros μ, σ^2 si tiene por densidad

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}$$

con $\mu \in \mathbb{R}$, $\sigma > 0$

Se denota como $X \sim N(\mu, \sigma^2)$

Así, se puede deducir de la definición anterior que los parametros: μ, σ^2, determinan completamente la distribución de X

Proposición: Sea $X \sim N(\mu, \sigma^2)$. Entonces:

1. $f(x)$ no negativa

2. $\displaystyle\int_{-\infty}^{+\infty} f(x)\,dx = 1$

$\boxed{D}$ 1.) Sabemos que $\sigma > 0$ y $e^x > 0$ $\forall x \in \mathbb{R}$

$\Rightarrow f(x) > 0 \quad \forall x \in \mathbb{R}$

2.) Si consideramos $y = \dfrac{x-\mu}{\sigma} \rightarrow dy = \dfrac{dx}{\sigma}$

Luego,
$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\,dy =$$

$$= \frac{1}{2\pi} \underbrace{\int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}}\,dy}_{\text{integral de Gauss}} = \frac{2\pi}{2\pi} = 1$$

Así, si $X \sim N(\mu, \sigma^2)$ entonces X es una v.a. continua.

3.1. Esperanza y varianza de $N(\mu, \sigma^2)$.

El siguiente resultado nos describe las principales medidas de dispersión de dicha variable:

<u>Proposición</u> - Sea $X \sim N(\mu, \sigma^2)$ entonces

1) $E[X] = \mu$

b) $E[X^2] = \mu^2 + \sigma^2$

c) $Var[X] = \sigma^2$

$\boxed{D}$ a) $E[X] = \int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \left[\begin{array}{l} x = \sigma y + \mu \\ \frac{x-\mu}{\sigma} = y \\ dy = \frac{1}{\sigma} dx \end{array} \right] =$ (A)

$$= \int_{-\infty}^{+\infty} (\sigma y + \mu) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy =$$

$$= \underbrace{\int_{-\infty}^{+\infty} \frac{\sigma}{\sqrt{2\pi}} y e^{-\frac{y^2}{2}} dy}_{0} + \frac{1}{\sqrt{2\pi}} \cdot \mu \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy =$$

$$= \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} \cdot \mu = \mu$$

b) $E[x^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{1}{\sigma} x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \overset{(*)}{=}$

$$\overset{(A)}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma y + \mu)^2 e^{-\frac{y^2}{2}} dy = \left[\begin{array}{l} \text{integrando por} \\ \text{partes y aplicar} \\ \text{lo anterior} \end{array} \right.$$

$$= \mu^2 + \sigma^2$$

c) $Var[X] = E[X^2] - EX^2 = E[(X-EX)^2]$

$$= \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

3.2. Tablas para la distribución normal

En primer lugar, cabe señalar que la función de distribución de $N(\mu, \sigma^2)$ viene dada por

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy$$

pero dicha integral no tiene primitiva conocida. Es decir, no existe ninguna expresión que use transformaciones de funciones elementales que represente $F(x)$. Así, antiguamente, para su cálculo acudíamos a tablas con estimaciones precisas del valor de la integral en un punto $x \in \mathbb{R}$. Actualmente, con la capacidad de cómputo actual se usan ordenadores o incluso calculadoras. Sin embargo, puesto que para cada par (μ, σ^2) tenemos una gráfica (distribución) diferente, necesitamos este resultado, antes de acudir a las tablas.

Proposición - Sea $X \sim N(\mu, \sigma^2)$ y $Z = \frac{X-\mu}{\sigma}$. Entonces, $Z \sim N(0,1)$ y

$$F_X(x) = F_Z\left(\frac{x-\mu}{\sigma}\right) \quad \forall x \in \mathbb{R}$$

[D] Observemos que si $X \sim N(\mu, \sigma^2) \Rightarrow$
$\Rightarrow E[X] = \mu$; $Var[X] = \sigma^2$

Luego,
$$E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{1}{\sigma} E[X-\mu] =$$

$$= \frac{1}{\sigma} E[X] - \frac{\mu}{\sigma} = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$$

$$Var[X] = \frac{1}{\sigma^2} Var[X-\mu] = \frac{1}{\sigma^2} Var[X] =$$

$$= \frac{\sigma^2}{\sigma^2} = 1$$

Además

$$F_Z(z) = P(Z \leq z) = P\left(\frac{X-\mu}{\sigma} \leq z\right) =$$

$$= P(X \leq \sigma z + \mu) = F_X(\sigma z + \mu)$$ $\square$

Así, a pesar que las tablas están calculadas para la $N(0,1)$ se puede hacer la transformación anterior. Por otro lado, puesto que la variable normal es ~~continua~~ simétrica, habitualmente sólo disponemos de la parte positiva. Veamos un ejemp

Ejemplo.- Sea $X \sim N(\mu, \sigma^2)$ y queremos calcular la probabilidad $P(X \leq -1)$.
Entonces, $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$.

$$P\left(\frac{X-\mu}{\sigma} \leq \frac{-1-\mu}{\sigma}\right) = P\left(Z \leq \frac{-1-\mu}{\sigma}\right)$$

Si $\mu = -1$ y $\sigma = 1$, entonces $P(Z \leq -2)$

Puesto que es simétrica

$$P(Z \leq -2) = P(Z \geq 2)$$

Luego, $P(Z \leq -2) = 1 - P(Z \leq 2)$

Así como $Z = 2$ si aparece en las tablas hemos terminado.

$\boxed{\text{TEOREMA}}$ (Aproximación de una binomial)

Sea $X \sim Bi(n,p)$ entonces

$$\lim_{n \to +\infty} P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq t \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx$$

Nota.- Así, si estandarizamos una $X \sim Bi(n,p)$ cuando $n >> 0$ entonces

$$\cancel{Bi}\ Z = \frac{X - np}{\sqrt{np(1-p)}} \sim N(0,1)$$

El Teorema anterior, se conoce como teorema de De Moivre - Laplace

3.3. Distribución de poblaciones normales

Supongamos que trabajamos sobre una población en la que estamos estudiando una v.a. continua $X \sim N(\mu, \sigma^2)$.

Si tomamos $(x_1, \dots, x_n)$ una muestra de tamaño n entonces los estadísticos más empleados son

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad ; \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \overline{x})^2$$

$\boxed{\text{TEOREMA}}$ Bajo las condiciones anteriores se cumple:

i) $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Lema de Fisher

ii) $\overline{X}$ y S^2 son independientes

iii) $E[S^2] = \sigma^2$

32' minutos 11.22'

Observación.- El teorema anterior es vital para la inferencia estadística y puesto que es el Tema siguiente consideramos que es la mejor manera de finalizar el tema

4. CONCLUSIONES