CIMAT Centro de Investigación en Matemáticas, A.C.

CONACYT

Curso Propedéutico: Probabilidad

MAEC-INEGI

07-Dic-2019

Unidad Monterrey

Temario

- Función de Distribución Acumulada
- Definición de media, mediana y varianza
- Función de generatriz de momentos
- Funciones de probabilidad discretas
  - Uniforme Discreta
  - Bernoulli
  - Binomial
  - Poisson
- Funciones de probabilidad continuas
  - Distribución Normal
- Distribuciones marginales
- Distribuciones condicionales

Función de Distribución Acumulada.

Dada una variable aleatoria X, su función de distribución acumulada $F(x)$ se define como:

$$F(x) = P(X \leq x) = \begin{cases} \sum_{t \leq x} f(t) & \text{caso discreto} \\ \int_{-\infty}^{x} f(t)dt & \text{caso continuo} \end{cases}$$

El subíndice en la sumatoria indica que se suman (acumulan) todos los valores de la función hasta el valor x, y los límites de la integral indican que se saca el área bajo la curva hasta llegar a x. En ambos casos se están sumando probabilidades a la izquierda de x. $F(x)$ tendrá un valor (que puede ser cero) para todos los números reales.

Con la distribución acumulada, puedes hacer los cálculos de,

- $P(X \leq a) = F(a)$
- $P(X > a) = 1 - P(X \leq a) = 1 - F(a)$
- $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$,

etc. Pero, debes tener cuidado, en el caso discreto importa si tienes $\leq$, o $<$, etc., pero no en el continuo:

- $P(X < a) = \begin{cases} F(a-\epsilon) & \text{si } X \text{ es discreta} \\ F(a) & \text{si } X \text{ es continua} \end{cases}$, donde ϵ es positivo pero cercano a cero

Definición: La media (valor esperado, esperanza matemática o promedio ponderado) de una v. a. X con distribución $f(x)$ se define como:

$$E(X) = \begin{cases} \sum_x x f(x) & \text{caso discreto} \\ \int_{-\infty}^{+\infty} x f(x) dx & \text{caso continuo} \end{cases}$$

Reflexionando en la fórmula para el caso discreto, podemos darnos cuenta que $E(X)$ representa un <u>promedio ponderado</u> ya que cada valor de X es "pesado" por su respectiva probabilidad. Para el ejemplo de la T. V. por cable

Teorema. Si X tiene distribución de probabilidad $f(x)$, entonces

$$E(g(X)) = \begin{cases} \sum_{x} g(x)f(x) & \text{caso discreto} \\ \int_{-\infty}^{+\infty} g(x)f(x)dx & \text{caso continuo} \end{cases}$$

Observa que en principio, para obtener el valor esperado de Y, necesitaríamos conocer $f(y)$, sin embargo, el teorema anterior nos evita el trabajo de calcular la distribución de Y. He aquí la importancia del resultado. Tenlo presente, pronto lo explotaremos.

Ejemplo. Según Y. Zimmels, los tamaños de partículas usadas en experimentos de sedimentación a menudo tienen distribuciones uniformes. Es importante estudiar la media y la varianza de los tamaños de las partículas, ya que en la sedimentación con mezclas de partículas de varios tamaños las partículas más grandes ocultan los movimientos de las más pequeñas. Suponiendo partículas esféricas con diámetros uniformemente distribuidos entre 0.01 y 0.05 centímetros. Encuentra la media de los *volúmenes* de estas partículas (recuerda que el volumen de una esfera es $\frac{4}{3}\pi r^3$).

Se sabe que la distribución de $X =$ diámetros de las partículas (en cms.), tiene una distribución uniforme (0.01,0.05), entonces

$$f(x) = \begin{cases} \frac{1}{.05-.01} & .01 \leq x \leq .05 \\ 0 & \text{todo lo demás} \end{cases}$$

$$= \begin{cases} 25 & .01 \leq x \leq .05 \\ 0 & \text{todo lo demás} \end{cases}$$

ahora, $R = \frac{X}{2}$, y como , $V =$ Volumen de partículas esféricas, entonces

$$V = \frac{4}{3}\pi\left(\frac{X}{2}\right)^3 = \frac{1}{6}\pi X^3$$

de donde,

$$E(V) = E(\tfrac{1}{6}\pi X^3) = \tfrac{\pi}{6}E(X^3)$$

por lo tanto,

$$\begin{aligned} E(V) &= \int_{-\infty}^{\infty}(\tfrac{\pi}{6}x^3)f(x)dx \\ &= \tfrac{\pi}{6}\int_{.01}^{.05}x^3(25)dx \\ &= 6.5\pi \times 10^{-6}\mathrm{cm}^3 \end{aligned}$$

Caso continuo.

Definición: **La mediana** es el valor de m que satisface:

$$P(X \leq m) = F(m) \geq \tfrac{1}{2}\,, \quad P(X \geq m) = 1 - P(X < m) \geq \tfrac{1}{2}$$

Observa que

$$P(X \leq m) = F(m) \geq \tfrac{1}{2}\,, \quad P(X \geq m) = 1 - P(X < m) \geq \tfrac{1}{2}$$

conducen a (si se despeja $P(X < m)$ de la segunda relación)

$$F(m) \geq \tfrac{1}{2}, \;\; P(X < m) \leq \tfrac{1}{2}$$

pero, en general

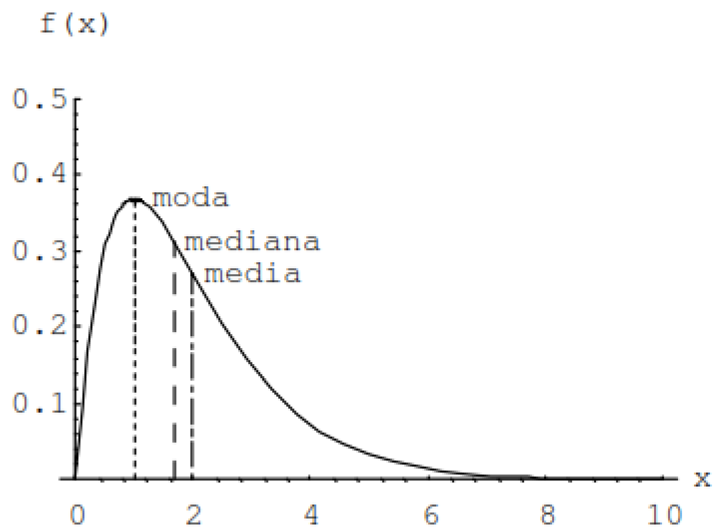
$$F(m) = P(X < m) + P(X = m)$$

entonces $P(X < m) = F(m) - P(X = m) \leq \frac{1}{2}$, de donde $F(m) \leq \frac{1}{2} + P(X = m)$.

$$\tfrac{1}{2} \leq F(m) \leq \tfrac{1}{2} + P(X = m)$$

y como en el caso continuo $P(X = x) = 0$, la relación se reduce a $\frac{1}{2} \leq F(m) \leq \frac{1}{2}$, esto es, a la izquierda de m y a su derecha (bajo la curva) hay un área de $\frac{1}{2}$.

En seguida te mostramos una distribución con sesgo positivo (cola larga a la derecha), en las que se señalan la moda, la media y la mediana. Observa cómo la media es "jalada" hacia las "colas largas" , y como la mediana no es tan sensible a estas colas. Repite el ejercicio para una distribución con sesgo negativo (cola larga a la izquierda).



Definición: La *varianza* de una v. a. X con distribución $f(x)$ es $E[(X-\mu)^2]$:

$$V(X) = E\left[(X-\mu)^2\right] = \begin{cases} \sum_x (x-\mu)^2 \cdot f(x)dx & \text{caso discreto} \\ \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot f(x)dx & \text{caso continuo} \end{cases}$$

Notación: $V(X) = \sigma^2 =$ Varianza de X.

Aunque pudimos utilizar $g(X) = |X-\mu|$, el valor absoluto siempre acarrea dificultades "técnicas". Inclusive, en vista de que las unidades de interés son las de X, se suele sacar la raíz cuadrada de la varianza: $\sqrt{V(X)} = \sqrt{\sigma^2} = \sigma$, y denominar a esta cantidad, la desviación estándar de X. Su significado (no su valor numérico) es el mismo que el de $V(X)$, pero en las unidades adecuadas.

La varianza y la desviación estándar nunca son negativas.

Cabe señalar que puede haber variables aleatorias con la misma media, pero distinta varianza. Recuerda que la varianza da el grado de dispersión de los valores de X (algo así como la concentración de los mismos alrededor de μ).

Tal como la media y mediana muestral, tenemos la varianza muestral para n datos

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2,$$

Propiedades del Valor Esperado y de la Varianza.

A continuación, veremos las propiedades más importantes del valor esperado de una función de una variable aleatoria, $g(X)$. Recuerda el teorema "poderoso" acerca del valor esperado de una función de X. Aquí, a y b, son constantes, y $g_1, g_2, g_3, h_1,$ y h_2 son funciones de X.

Si $g_1(X) = a,\ g_2(X) = aX + b,$ y $g_3(X) = ah_1(X) + h_2(X)$, entonces

$$E(g_1(X)) = E(a) = a$$

$$E(g_2(X)) = E(aX + b) = aE(X) + b = a\mu + b$$

$$E(g_3(X)) = E(ah_1(X) + h_2(X)) = aE(h_1(X)) + E(h_2(X)).$$

Estas propiedades se pueden demostrar, ya sea para el caso discreto usando sumatorias o el caso continuo usando integrales. Completa las demostraciones de esta manera.

La primera propiedad se puede interpretar pensando en que el promedio de una variable que siempre toma un único valor es ese mismo valor. En la segunda propiedad, si se multiplica la variable por una constante, ésta resulta "escalada", se le cambia su escala; y al sumarle (restarle) una constante, es "trasladada", de tal forma que su promedio será escalado y trasladado en igual forma.

Ejercicio: Si el modelo $Y = \mu + \epsilon$, lo escribimos como $\epsilon = Y - \mu$, ¿cuál es el valor esperado de ϵ?

MOMENTOS DE UNA V. A. X

Notemos que, $E(X)$, y $E((X-\mu)^2)$ nos dan información sobre el problema que se estudia ¿Nos darán también información $E((X-\mu)^3)$ y $E((X-\mu)^4)$? La respuesta es, sí.

La cantidad $\gamma = E((\frac{X-\mu}{\sigma})^3) = \frac{E((X-\mu)^3)}{\sigma^3}$ se conoce como: $coeficiente\,de\,asimetría$ y nos da el grado de simetría de la distribución $f(x)$ alrededor de la media .

- Si γ es aproximadamente cero, habrá simetría alrededor de la media,
- Si $\gamma > 0$ será sesgada a la derecha (cola larga a la derecha),
- Si $\gamma < 0$ será sesgada a la izquierda (cola larga a la izquierda).

La cantidad $\delta = E((\frac{X-\mu}{\sigma})^4) - 3 = \frac{E((X-\mu)^4)}{\sigma^4} - 3$, se llama $coeficiente\,de$ *curtosis* y nos dice el grado de semejanza que se tiene con respecto a una curva tipo campana (que tanta semejanza hay con la distribución normal). Si δ es aproximadamente cero la $f(x)$ se *parece*, a una densidad normal.

Algunas gráficas con el valor de sus coeficientes de asimetría y curtosis son mostradas, también se muestra una curva normal (campana). Observa el significado de "sesgo a la derecha" y "sesgo a la izquierda".

Existen los coeficientes análogos para un conjunto de datos experimentales y su interpretación es similar

- $\widehat{\gamma} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^3}{ns^3}$ (Coeficiente muestral de Asimetría)

- $\widehat{\delta} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^4}{ns^4} - 3$ (Coeficiente muestral de Curtosis).

Considerando todos los binomios en las fórmulas para δ y γ, a las potencias respectivas, aplicando propiedades del valor esperado, se obtienen expresiones como: $E(X^2), E(X^3)$ y $E(X^4)$. Estas cantidades son los **momentos** de la v.a. X (respecto del origen).

Definición: Los *momentos* de orden r de la v. a. X (centrada en el origen) se definen como el valor esperado de $g(X) = X^r$, con r, un entero positivo:

$$\mu_r' = E(X^r) = \begin{cases} \sum_x x^r f(x) & \text{caso discreto} \\ \int_{-\infty}^{+\infty} x^r f(x) dx & \text{caso continuo} \end{cases} \qquad \text{para } r \text{ un entero positivo}$$

Notación: $E(g(X)) = E(X^r) = \mu_r'$

La función *GENERATRIZ DE MOMENTOS*

Existe otra función de X cuyo valor esperado es de gran utilidad: $g(X) = e^{tX}$; a este valor esperado se le llama, la función generatriz (generadora) de momentos de X, y se denota en forma especial como $M_X(t)$, donde t es un parámetro real definido sobre un intervalo que contenga al cero, veamos porqué se llama así,

Definición: La generatriz de momentos $M_X(t)$ de una v.a. $.X$, si existe, es

$$M_X(t) = E\left(e^{tX}\right) = \begin{cases} \sum\limits_x e^{tx} f(x) & \text{caso discreto} \\ \int\limits_{-\infty}^{+\infty} e^{tx} f(x)dx & \text{caso continuo} \end{cases}$$

Del Cálculo se sabe que la serie de potencias de : e^u alrededor del cero, está dada por

$$e^u = 1 + u + \tfrac{u^2}{2!} + \tfrac{u^3}{3!} + \cdots$$ entonces, si tomamos $u = tX$, obtenemos

$$e^{tX} = 1 + tX + \tfrac{(tX)^2}{2!} + \tfrac{(tX)^3}{3!} + \cdots$$ y al tomar su valor esperado nos queda

$$\begin{aligned} E\left(e^{tX}\right) &= E\left[1 + tX + \tfrac{(tX)^2}{2!} + \tfrac{(tX)^3}{3!} + \cdots\right] \\ &= 1 + E(tX) + E\left[\tfrac{(tX)^2}{2!}\right] + E\left[\tfrac{(tX)^3}{3!}\right] + \cdots \\ &= 1 + tE(X) + \tfrac{t^2}{2!}E(X^2) + \tfrac{t^3}{3!}\,E(X^3) + \cdots \end{aligned}$$

Si ahora derivamos con respecto a la variable real t, y evaluamos en $t = 0$:

$$M'_X(t) = E(X) \; + \; \tfrac{2t}{2!}\, E(X^2) \; + \; \tfrac{3t^2}{3!}\, E(X^3) \; + \cdots$$

entonces,

$$M'_X(0) = E(X)$$

también,

$$M''_X(t) = E(X^2) \; + \; \tfrac{2t}{2!}\, E(X^3) \; + \; \tfrac{3t^2}{3!}\, E\big(X^4\big) + \cdots$$

de donde,

$$M''_X(0) = E(X^2)$$

y en general

$$M_X^{(n)}(0) = E(X^n) \qquad\qquad \forall \;\; n = 1, 2, 3, \ldots$$

Así, si tenemos la $M_X(t)$ de una v. a., sólo tenemos que derivarla un número r de veces y evaluarla en $t = 0$, para obtener el r-ésimo momento de X.

<u>Ejemplo</u>. Calcular $M_X(t)$, para la v. a. X, cuya distribución de probabilidad es $f(x) = \frac{1}{2}e^{-\frac{1}{2}x}$, $x > 0$.

Propiedades de la Generatriz de Momentos

Sea una v. a. con f. g. m. $M_X(t)$, entonces las generatrices de momentos de las variables aleatorias definidas como $Y = \frac{X}{b}$, $Z = \frac{X-a}{b}$, son, respectivamente:

- $M_Y(t) = M_{\frac{X}{b}}(t) = E(e^{t\frac{X}{b}}) = E(e^{(\frac{t}{b})X}) = E(e^{(t^*)X}) = M_X(t^*) = \boldsymbol{M_X(\frac{t}{b})},$ donde $t^* = \frac{t}{b}$

- $$
\begin{aligned}
M_Z(t) = M_{\frac{X-a}{b}}(t) = E(e^{t(\frac{X-a}{b})}) = E(e^{(\frac{t}{b})X-\frac{a}{b}t}) = E(e^{(\frac{t}{b})X}e^{-\frac{a}{b}t}) &= e^{-\frac{a}{b}t}E(e^{(\frac{t}{b})X}) \\
&= e^{-\frac{a}{b}t}E(e^{(t^*)X}) \\
&= e^{-\frac{a}{b}t}M_X(t^*) \\
&= \boldsymbol{e^{-\frac{a}{b}t}M_X(\frac{t}{b})}
\end{aligned}
$$

Observa que $M_{\frac{X-a}{b}}(t)$ puede escribirse como $M_{\frac{X}{b}-\frac{a}{b}}(t)$, así, la segunda propiedad queda como

$$M_Z(t) = M_{\frac{X}{b}-\frac{a}{b}}(t) = e^{-\frac{a}{b}t}M_X(\tfrac{t}{b})$$

1. - Distribución Uniforme Discreta.

Cuando se tiene un número **finito** de resultados de un experimento: $x_1, x_2, \ldots, x_n$, cada uno de ellos **igualmente probable**, se dice que se tiene una variable aleatoria que puede ser modelada por una distribución uniforme discreta . Notación:

$$X \sim U(x_1, x_2, \ldots, x_n)$$

Se lee: X se distribuye como una variable aleatoria uniforme.

Este tipo de modelo aparece típicamente en la selección de muestras al azar.

El único parámetro de la distribución es: n, el número posible de resultados.

Como la suma de todas las probabilidades debe ser uno, entonces

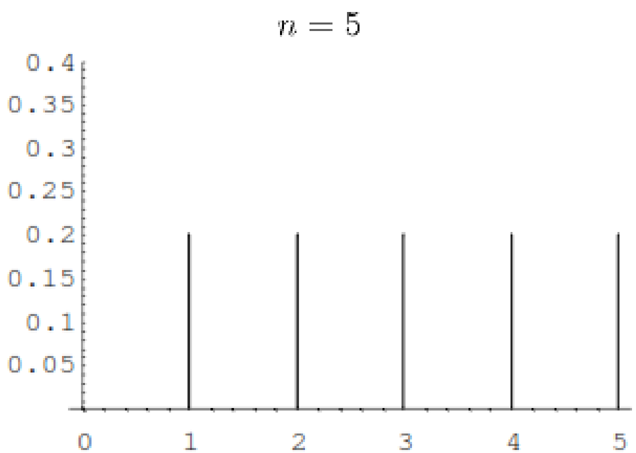
$$f(x) = \begin{cases} \frac{1}{n} & \text{para } x = x_1, x_2, \ldots, x_n \\ 0 & \text{otro caso} \end{cases}$$

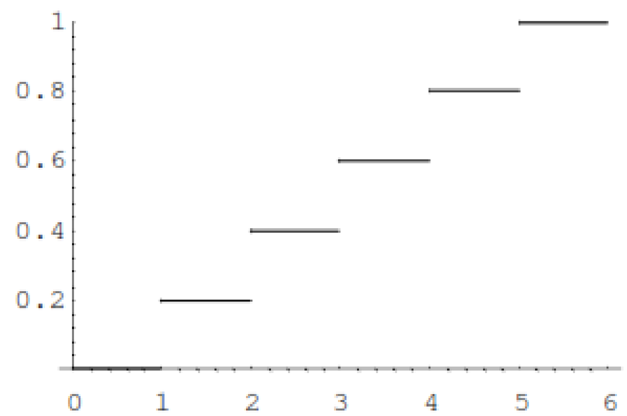
Bajo esta definición , es claro que: $f(x) \geq 0$, para toda x, y $\sum_{x=x_i}^{x_n} f(x) = n(\frac{1}{n}) = 1$.

El caso más común de esta distribución es cuando la variable toma sólo valores enteros,

$$X = 1, 2, 3, \ldots, n.$$

El resto de los resultados los estableceremos para estos valores de la variable aleatoria. El parámetro (valor que identifica unívocamente al modelo) de la distribución es n, el número total de objetos. Se dice en este caso que se trata de un espacio de probabilidad equiprobable.





$$E(X) = \sum_{x=1}^{n} x\frac{1}{n} = \frac{1+2+\cdots+n}{n} = \boldsymbol{\frac{n+1}{2}}$$

y

$$\begin{aligned} V(X) &= E(X^2) - [E(X)]^2 \\ &= \sum_{x=1}^{n} x^2\frac{1}{n} - \left(\frac{n+1}{2}\right)^2 \\ &= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{2n^2+3n+1}{6} - \frac{n^2+2n+1}{4} \\ &= \frac{2(2n^2+3n+1)-3(n^2+2n+1)}{12} = \frac{4n^2+6n+2-3n^2-6n-3}{12} \\ &= \boldsymbol{\frac{n^2-1}{12}} \end{aligned}$$

<u>Función Generatriz de Momentos</u>:

$$\begin{aligned} M_X(t) &= E\left(e^{tX}\right) = \sum_{x=1}^{n} e^{tx}\cdot f(x) = \sum_{x=1}^{n} e^{tx}\cdot\frac{1}{n} \\ &= \frac{1}{n}\sum_{x=1}^{n}\left(e^t\right)^x = \frac{1}{n}\left[\frac{e^t-(e^t)^{n+1}}{1-e^t}\right] = \boldsymbol{\frac{1}{n}\cdot\frac{e^t(1-e^{nt})}{1-e^t}}, \quad \forall t \end{aligned}$$

2.- Distribución Bernoulli

Definamos un experimento en el cual hay únicamente dos posibles resultados. A uno de ellos le llamamos "éxito" al otro "fracaso". Este tipo de variables aparece frecuentemente en nuestros conteos, por ejemplo, si pensamos en clasificar nuestros productos como: defectuoso, no defectuoso, grande o pequeño, azul y blanco, sí o no, etc.

Generalmente le asignamos un valor de 1 al "éxito" y un valor de 0 al "fracaso". Notemos que la asignación de los valores numéricos es arbitraria. Este procedimiento sirve de base para la construcción de otras distribuciones de gran utilidad.

Definamos $X =$ # de "éxitos", entonces

$$x = \begin{cases} 1 & \text{si } X = \text{"éxito", esto con probabilidad } p \\ 0 & \text{si } X = \text{"fracaso", esto con probabilidad } (1-p) = q \end{cases}$$

o bien,

$$f(x) = p^x (1-p)^{1-x}, \quad \text{para } x = 0, 1 \qquad (\textbf{Distribución de Bernoulli})$$

Notación: $X \sim Bernoulli\,(p)$.
El único parámetro de la distribución de esta variable aleatoria es p, la probabilidad de éxito.

El experimento que guió al modelo, dos posibles resultados con probabilidades p y q $(p+q=1)$, se denomina experimento Bernoulli.

Media y Varianza:

$$E(X)=\sum_{x=0}^{1} x\cdot f(x)=0\cdot(1-p)+1\cdot p=\boldsymbol{p}$$

$$E(X^2)=\sum_{x=0}^{1} x^2\cdot f(x)=0^2\cdot(1-p)+1^2\cdot p=p$$

$$V(X)=E(X^2)-[E(X)]^2=p-p^2=p\underbrace{(1-p)}=\boldsymbol{pq}$$

Función Generatriz de Momentos:

$$M_X(t)=E\left(e^{tX}\right)=\sum_{x=0}^{1} e^{tx}\cdot f(x)=e^{t\cdot 0}(1-p)+pe^{t\cdot 1}$$

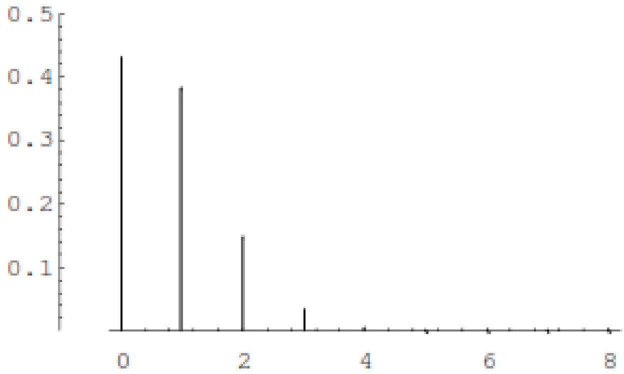
$$=(1-p)+pe^t=\boldsymbol{q+pe^t},\quad \forall t$$

$$f(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & x = 0, 1, \ldots, n \\ 0 & \text{otro caso} \end{cases} \qquad \textbf{(Distribución Binomial)}$$

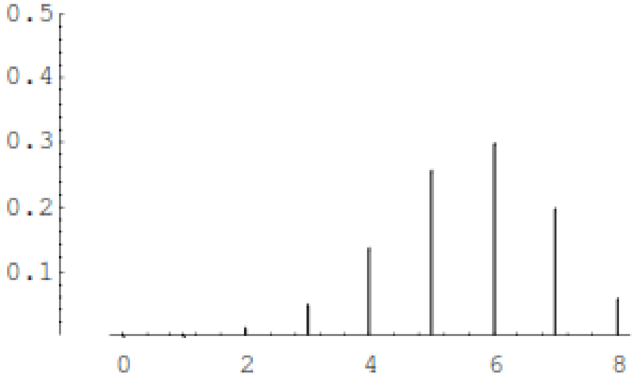
En este caso se dice que X sigue una distribución binomial con parámetros (n, p).

Notación: $X \sim B(n, p)$.

$n = 8.p = 0.1$



$n = 8, p = 0.7$



Su Función Generatriz de Momentos es:

$$M_X(t) = E\left(e^{tX}\right) = \sum_{x=0}^{n} e^{tx} \cdot f(x) = \sum_{x=0}^{n} e^{tx} \cdot \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \underset{\substack{\downarrow \\ (e^t)^x}}{e^{tx}} \cdot \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} \left(e^t p\right)^x q^{n-x}$$

De acuerdo al binomio de Newton con $a = q$, $b = e^t p$:

$$M_X(t) = E\left(e^{tX}\right) = \boldsymbol{(q + pe^t)^n}$$

Media y Varianza

Utilizaremos la generatriz de momentos. Sabemos que $E(X) = M_X'(0)$ y $E(X^2) = M_X''(0)$ entonces

$$M_X'(t) = n(q + pe^t)^{n-1}(pe^t)$$
$$M_X'(0) = n(q + pe^0)^{n-1}(pe^0)$$

puesto que $e^0 = 1$ y $p + q = 1$ tenemos:

$$M_X'(0) = n(1)^{n-1} p$$

por lo tanto

$$E(X) = \boldsymbol{np}$$

ahora,

$$M_X''(t) = n(q+pe^t)^{n-1}pe^t + pe^t\left[n(n-1)(q+pe^t)^{n-2}pe^t\right]$$

$$E(X^2) = M_X''(0) = np + n(n-1)p^2$$

entonces

$$\begin{aligned} V(X) = E(X^2) - [E(X)]^2 &= np + n(n-1)p^2 - (np)^2 \\ &= np + n^2p^2 - np^2 - n^2p^2 \\ &= np - np^2 \\ &= np(1-p) \end{aligned}$$

$$V(X) = \boldsymbol{npq}$$

La variable aleatoria Binomial como una suma de variables aleatorias Bernoulli

Hemos mencionado que cada X_i, es una v. a. Bernoulli y que estas son independientes. Considera la variable aleatoria Y definida como la suma de n variables aleatorias Bernoulli (p), X_i, esto es, como la suma de ceros y unos producidos por los posibles resultados de las variables Bernoulli, podemos imaginarnos lo anterior como realizando n repeticiones de un experimento Bernoulli en el que cada resultado será independiente de los otros, un **muestreo aleatorio**.

$$Y = \sum_{i=1}^{n} X_i, \quad \text{con } X_i \sim Bernoulli\,(p) \text{ e independientes entre sí.}$$

Su valor esperado, recordando las propiedades del valor esperado, es

$$E(Y) = E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = \boldsymbol{np},$$

su varianza

$$V(Y) = V(\sum_{i=1}^{n} X_i),$$

veremos posteriormente que la varianza de una suma de variables aleatorias independientes es la suma de las varianzas de cada v. a. individual (multiplicada por el cuadrado de su coeficiente respectivo), entonces

$$V(Y) = \sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} pq = \boldsymbol{npq}$$

vemos que estos resultados coinciden con los de la v. a. Binomial (n, p), pero no es suficiente para afirmar que $Y \sim$ Binomial (n, p); si su generatriz es la misma que la de una v. a. binomial podremos afirmar que Y es binomial:

$$M_Y(t) = E(e^{tY}) = E(e^{t\sum_{i=1}^{n} X_i}) = E(e^{t(X_1+X_2+\cdots+X_n)}) = E(e^{tX_1}e^{tX_2}\cdots e^{tX_n})$$

de nuevo, como las variables aleatorias X_i son independientes, el último término se puede escribir como

$$M_Y(t) = E(e^{tX_1})E(e^{tX_2})\cdots E(e^{tX_n}) = (q+pe^t)(q+pe^t)\cdots(q+pe^t) = (q+pe^t)^n$$

que es idéntica a la generatriz de momentos, por lo tanto nuestra v. a. Y, producto de un muestreo determinado, es una v. a. Binomial (n, p)., ver teorema de la generatriz en conceptos básicos. ¿Qué se obtiene si en la sumatoria que define Y usamos $n = 1$?

Es muy importante recalcar el hecho de que cada posible resultado del experimento es una v. a. con las mismas propiedades: misma distribución, mismo parámetro, y que el resultado de cada una de ellas será independiente del resultado de las otras. Estas condiciones se resumen diciendo que se realizó un Muestreo Aleatorio.

5.- Distribución Poisson.

Supóngase ahora que estamos interesado en el número de "éxitos", pero ahora preguntándonos por la posibilidad de que ocurran en un cierto intervalo de tiempo o espacio, por ejemplo

- Número de defectos en una plancha de acero de $1m^2$
- Número de bacterias en $1\,cm^3$ de agua potable
- Número de entregas de materia prima entre las 8:00 A. M. y las 1.00 P. M.
- Número de llamadas que llegan a un conmutador en un minuto.

En estos casos interesa la variable aleatoria

$X =$ # de "éxitos" obtenidos en un cierto intervalo (temporal o espacial).

Nuestros supuestos son:

- El número promedio de "éxitos", λ, sobre el intervalo dado se mantiene constante.

Se considera que el comportamiento es estable. Incluso, si se estuviera interesado en un múltiplo t (no necesariamente entero) del intervalo original, el valor λ_1, correspondiente será $\lambda_1 = \lambda t$.

- Los éxitos aparecen en forma aleatoria en intervalos de la misma magnitud, los intervalos no se traslapan, y son independientes entre ellos.

Se considera que si en un sub-intervalo del intervalo dado ocurre un determinado número de "éxitos", no implica que en otro sub-intervalo de la misma amplitud tendrá ese mismo número, mayor que él, o menor. Sólo se espera que en promedio se tenga el mismo número.

Observa que no se hace un número de pruebas hasta obtener "éxitos", sino que se cuenta el número de éxitos obtenidos en un continuo espacial o temporal y por ende, cuando el intervalo tiende a cero la probabilidad de éxito también tiende a cero.

- El número posible de "éxitos" es infinito.

Supongamos que tenemos un proceso Poisson y lo observamos en un cierto intervalo, digamos de una unidad, y asumimos que λ es el número promedio de éxitos en esa unidad observada.

Por otra parte, consideremos esa misma unidad y que:

Se realizan 10 observaciones y se cuenta el número de "éxitos", cada uno con probabilidad p, obtenidos en 10 subdivisiones. Es natural pensar que hay una probabilidad relativamente pequeña de observar más de un éxito, en cada subdivisión.
Se realizan 1000 observaciones y se cuenta el número de "éxitos", cada uno con probabilidad p (más pequeña que en el anterior caso), obtenidos en mil subdivisiones del intervalo de tal manera que p será tan pequeña que hay una probabilidad de cero de observar más de un éxito, en cada subdivisión.

El número de "éxitos" Y en n pruebas independientes con probabilidad p, se puede modelar ahora como una v. a. Binomial, cuya media es $E(Y) = np$, que debe coincidir con el número promedio de éxitos λ en el intervalo considerado.
Notemos que $\lambda = np \;\; implica\, que\, p = \frac{\lambda}{n}$ y como λ es fija, p será más y más pequeña a medida que n crece.

Esto nos lleva a establecer el siguiente resultado:

La distribución de Y es $Poisson$ con $\lambda = np$ constante, cuando el número de observaciones tiende a infinito en el intervalo, con p siendo más pequeña conforme n aumenta.

Demostración:

Si $X \sim B(n,p)$, entonces

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad np = \lambda \ \rightarrow \ p = \frac{\lambda}{n}$$

$$f(x) = \frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^{x}\left(1-\frac{\lambda}{n}\right)^{n-x} = \frac{n\,(n-1)\,(n-2)\,\ldots\,(n-(x-1))\,(n-x)!}{x!\,(n-x)!}\,\frac{\lambda^x}{n^x}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\frac{n}{n}\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\ldots\left(1-\frac{(x-1)}{n}\right)}{x!}\lambda^x \left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$f(x) = \frac{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{(x-1)}{n}\right)}{x!}\ \lambda^x \ \left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda}\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{(x-1)}{n}\right) \rightarrow 1 \text{ cuando } n \rightarrow +\infty$$

$$\left(1-\frac{\lambda}{n}\right)^{-x} \rightarrow \ 1 \ \text{ cuando } n \rightarrow +\infty$$

$$x! \quad \& \quad \lambda^x \quad \text{quedan igual}$$

El único término faltante es $\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}$

$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}$ tiene la forma $1^{-\infty}$ que es una forma indeterminada.

Sea $y=\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}$ entonces $ln\, y = -\frac{n}{\lambda}\, ln\left(1-\frac{\lambda}{n}\right)=\frac{-ln\left(1-\frac{\lambda}{n}\right)}{\left(\frac{\lambda}{n}\right)}$

$$\lim_{n\to\infty}(lny)=\lim_{n\to\infty}\left[\frac{-ln\left(1-\frac{\lambda}{n}\right)}{\frac{\lambda}{n}}\right]$$ forma indeterminada del tipo $\frac{0}{0}$

aplicando la regla de L'Hopital

$$\lim_{n\to\infty}-\frac{\frac{1}{\left(1-\frac{\lambda}{n}\right)}\left(\frac{\lambda}{n^2}\right)}{\left(-\frac{\lambda}{n^2}\right)}=\lim_{n\to\infty}\frac{1}{\left(1-\frac{\lambda}{n}\right)}=1$$

$$\therefore \lim_{n\to\infty}(lny)=1$$

Cuando f es una función continua
$$f\left(\lim_{x\to a}y\right)=\lim_{x\to a}f(y)$$

$ln\left(\lim_{n\to\infty}y\right)=1$ por ser la función logaritmo natural una función continua

$\therefore \lim_{n\to\infty}y=e \qquad \therefore \qquad \left(1-\frac{\lambda}{n}\right)^{-n/\lambda}\to e$ cuando $n\to\infty$

Usando todos estos resultados obtenemos la función de probabilidad

$$f(x)=\frac{\lambda^x e^{-\lambda}}{x!} \qquad \text{para } x=0,1,2,\ldots$$

(Distribución Poisson)

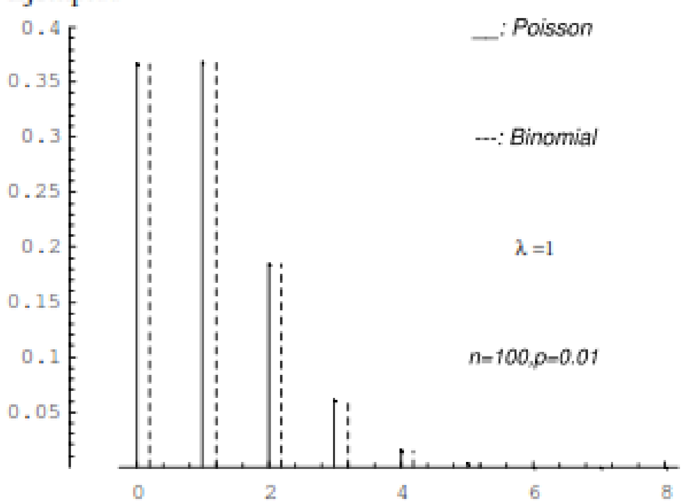
CIMAT

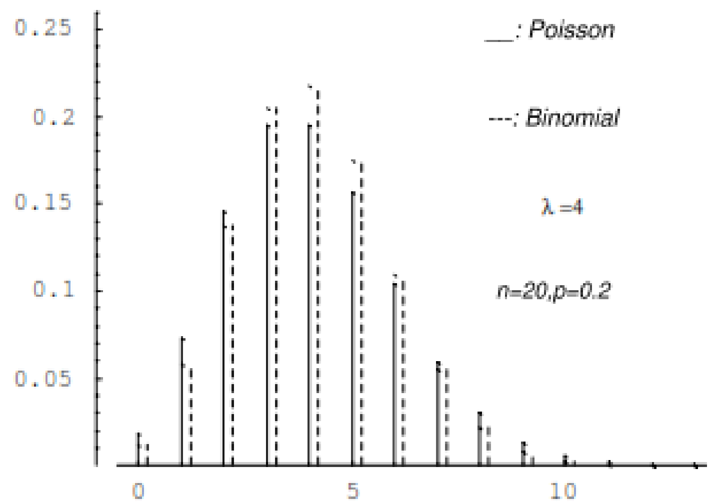
1) La fórmula para asignar probabilidades, como siempre, es válida bajo una serie de supuestos que debemos verificar se cumplan razonablemente en nuestros estudios particulares (los marcados con $\bullet$) . Por otro lado, es el resultado de un proceso matemático no necesariamente intuitivo, como lo había sido en los casos anteriores.

2) Aunque no siempre se tiene que pensar que un modelo Poisson proviene de una variable binomial tomada en circunstancias límite como se mostró arriba, sí nos deja una herramienta de aproximación muy útil. Más aún, la aproximación es utilizada a favor nuestro en el sentido de que, los cálculos bajo una Poisson son, en general, más simples que con una binomial.

La variable aleatoria Poisson se puede usar para calcular probabilidades binomiales en forma aproximada cuando n es muy grande y p es pequeña. La aproximación es buena cuando $n \geq 20$ y $p \leq 0.05$, cuando $n \geq 100$ y $np \leq 10$ la aproximación es excelente.

Ejemplos





Función generatriz de momentos:

$$M_X(t) = E(e^{tx}) = \sum_{x=0}^{+\infty} e^{tx} f(x) = \sum_{x=0}^{+\infty} e^{tx} \, \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{+\infty} \frac{e^{tx}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{+\infty} \frac{(e^t\lambda)^x}{x!}$$ Recordando el desarrollo en series de potencias para la función

exponencial: $e^z = 1 + z + \frac{z^2}{2} + \frac{z^3}{3!} + \frac{z^4}{4!} + \ldots$ obtenemos

$$M_X(t) = e^{-\lambda} e^{\lambda e^t}$$

$$M_X(t) = \boldsymbol{e^{\lambda(e^t - 1)}}$$

Ahora, utilizando la relación entre $M_X(t)$ y $E(X^r)$, obtenemos la media y la varianza

$$E(X) = \mu = \mu_1' = M_X'(t) \mid_{t=0}$$

$$M_X'(t) = e^{\lambda(e^t-1)} \cdot \lambda e^t \qquad\qquad M_X'(0) = e^{\lambda(e^0-1)} \lambda e^0 = \lambda$$

Por lo tanto $E(X) = \boldsymbol{\lambda}$. Ahora

$E(X^2) = \mu_2' = M_X''(t) \mid_{t=0}$

$$M_X''(t) = e^{\lambda(e^t-1)}\lambda e^t + \lambda e^t e^{\lambda(e^t-1)}\lambda e^t \qquad M_X''(0) = \lambda + \lambda^2$$

Por lo tanto $\quad E(X^2) = \lambda + \lambda^2.$

Entonces $\quad V(X) = E(X^2) - [E(X)]^2$

$$= \lambda + \lambda^2 - \lambda^2 = \boldsymbol{\lambda}$$

Ejemplo. ¿Qué tan grande debe ser el radio de una región circular de muestreo en un bosque para que la probabilidad de encontrar al menos un chapulín sea de 0.99? Sabiendo que se esperan encontrar 2 chapulines por m^2, según las condiciones de la región.

Definición. Una v. a. Z se dice que tiene una distribución **Normal Estándar** si su f. d. p. está dada por

$$f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2}, \qquad -\infty < Z < \infty.$$

- Es simétrica alrededor de un eje vertical que pasa por $z = 0$.
- Su mediana coincide con su media, es decir $m = \mu = 0$
- Su varianza es 1 y tiene puntos de inflexión en $\pm\, 1$
- Su valor máximo ocurre en $z = 0$ y es $\frac{1}{\sqrt{2\pi}}$
- $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2} dz = 1$
- La integral $\int_a^b \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2} dz$, ($= P(a \leq Z \leq b)$) con a y b números reales, no puede evaluarse analíticamente, esto es, no tiene una antiderivada, sólo se puede hacer mediante métodos numéricos, ya en épocas de DeMöivre se calculaba de esa manera. Nosotros usaremos las tablas respectivas.

- La acumulada $F_Z(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}u^2} du$, y se denota como $\Phi(z)$.

Media y Varianza.

El valor esperado de Z es

$$\mu = E(Z) = \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2} dz = 0$$

ya que el integrando es una función impar (si lo deseas puedes integrar mediante un cambio de variable).

La varianza es

$$\sigma^2 = V(Z) = E\left[(z-\mu)^2\right] = E\left[z^2\right] = \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2} dz$$

$$= 1$$

la cual se obtiene integrando por partes dos veces.

Entonces la media y la varianza son cero y uno respectivamente, esto se escribe abreviadamente como:
Notación: $Z \sim N(0, 1)$, en palabras, Z tiene una distribución normal estándar con media cero y varianza uno. El uso de la letra Z para una v. a. normal estándar se ha generalizado en la literatura.

Función Generatriz de Momentos.

$$M_Z(t) = E(e^{tz}) = \int_{-\infty}^{\infty} e^{tz} f(z)\, dz = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2} dz$$

Este es un buen momento para hacerte una recomendación respecto del cáculo de integrales que se ven complicadas; trata de identificar la integral que estés resolviendo con alguna que hayas manejado anteriormente, haciendo las analogías necesarias puedes obtener su valor, por ejemplo, teníamos que

$$M_Z(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2 + tz} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2tz)} dz$$

ompletando cuadrados y separando

$$\begin{aligned} M_Z(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[z^2 + 2tz + (\frac{2t}{2})^2 - (\frac{2t}{2})^2\right]} dz \\ &= e^{\frac{1}{2}t^2} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z+t)^2} dz\right] \\ &= e^{\frac{1}{2}t^2} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(u)^2} du\right] \\ &= e^{\frac{1}{2}t^2} \left[1\right] = e^{\frac{1}{2}t^2} \end{aligned}$$

De hecho, la distribución de $X = Z\sigma + \mu$ es

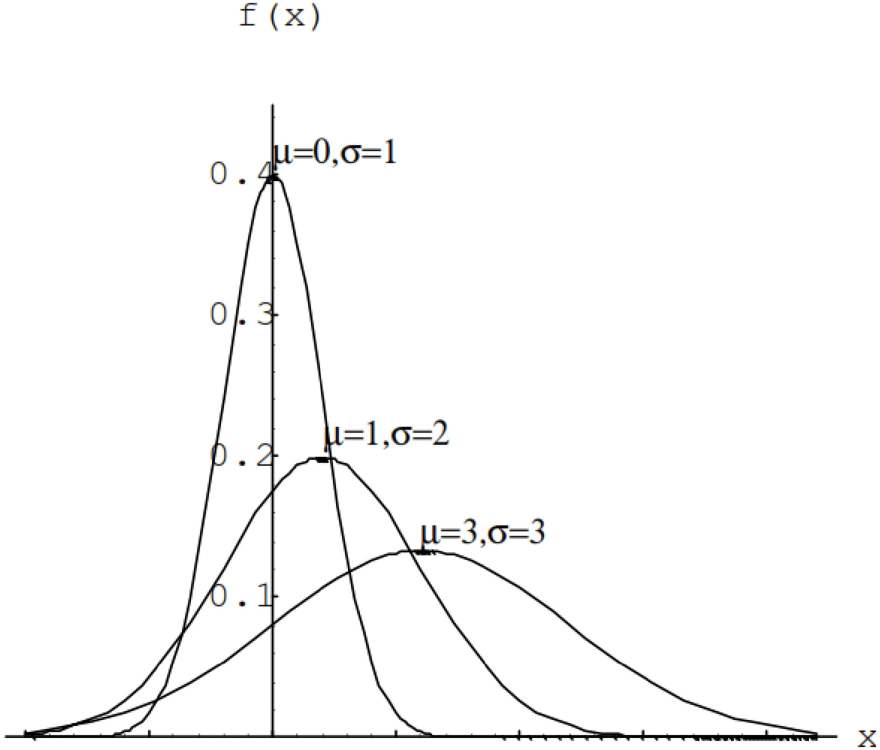
$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}$$

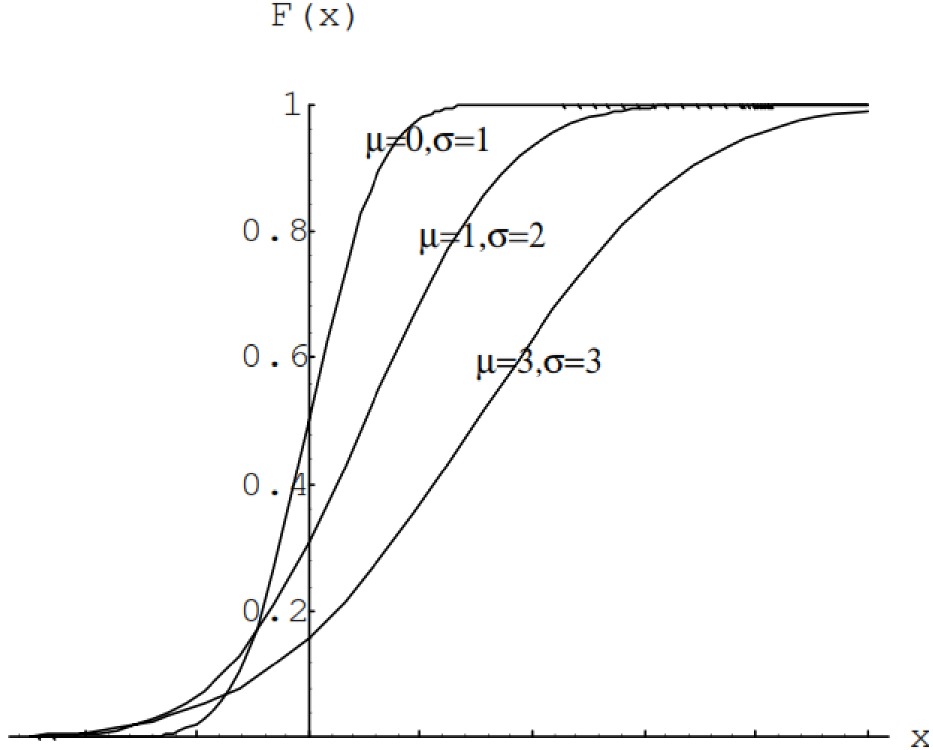
Con media μ y varianza σ^2, y se dice que se distribuye normal con parámetros μ y σ.Notación: $X \sim N(\mu, \sigma^2)$. La generatriz de momentos

$$M_X(t) = M_{Z\sigma+\mu}(t) = e^{\mu t} M_Z(t\sigma) = e^{\mu t} e^{\frac{1}{2}(t\,\sigma)^2}$$

$$= \boldsymbol{e^{\mu t + \frac{1}{2} t^2 \sigma^2}}$$

donde se usó la propiedad, demostrada en conceptos básicos,


f(x)
0.4
0.3
0.2
0.1
x
μ=0,σ=1
μ=1,σ=2
μ=3,σ=3


F(x)
1
0.8
0.6
0.4
0.2
x
μ=0,σ=1
μ=1,σ=2
μ=3,σ=3

Ejemplo. Se tiene un proceso industrial donde se producen baleros y se toma la medición de su diámetro. La especificación del producto toma como bueno un balero con diámetro en el rango 3.0 ± 0.01 cm. Si la variable : diámetro, se puede considerar normal con media $\mu = 3$ y $\sigma = 0.005$

a) ¿Qué proporción de baleros caen dentro de especificación?

b) Se realiza una investigación en la que es necesario revisar el 5% de los baleros producidos que se encuentran a la izquierda del límite de especificación derecho, esto es, el 5% de los producidos a la izquierda de 3.01. ¿A partir de qué valor de la v. a. X se efectuaría esta revisión?

Distribuciones Marginales

Definition 3.23 Si (X,Y) tiene distribución conjunta con función de masa $f_{X,Y}$

Entonces la distribución marginal para X es

$$f_X(x) = \mathbb{P}(X = x) = \sum_y \mathbb{P}(X = x, Y = y) = \sum_y f(x, y) \qquad (3.3)$$

Y la marginal para Y se define como

$$f_Y(y) = \mathbb{P}(Y = y) = \sum_x \mathbb{P}(X = x, Y = y) = \sum_x f(x, y). \qquad (3.4)$$

Distribuciones marginales

Definition 3.25 Para variables aleatorias continuas las marginales son

son

$$f_X(x) = \int f(x,y)dy, \quad \text{y} \qquad f_Y(y) = \int f(x,y)dx. \qquad (3.5)$$

Las distribuciones marginales correspondientes se denotan por Fx y Fy

Ejemplo Suponga que

$$f(x,y) = \begin{cases} x+y & \text{si} \quad \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & \text{En otro lado} \end{cases}$$

Entonces

$$f_Y(y) = \int_0^1 (x+y)\,dx = \int_0^1 x\,dx + \int_0^1 y\,dy = \frac{1}{2} + y. \quad \blacksquare$$

Distribuciones condicionales

La función de distribución condicional es

$$f_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

$if\ f_Y(y) > 0.$

En el caso de variables aleatorias continuas

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

Asumiendo que $f_Y(y) > 0.$ Entonces,

$$\mathbb{P}(X \in A|Y = y) = \int_A f_{X|Y}(x|y)dx.$$

Ejemplo Sea

$$f(x,y) = \begin{cases} x+y & \text{if } 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{En otro lado} \end{cases}$$

Encontremos $\mathbb{P}(X < 1/4 | Y = 1/3)$. Anteriormente vimos que $f_Y(y) = y + (1/2)$. Por tanto,

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{x+y}{y+\frac{1}{2}}.$$

Entonces,

$$\begin{aligned} P\left(X < \frac{1}{4} \,\middle|\, Y = \frac{1}{3}\right) &= \int_0^{1/4} f_{X|Y}\left(x \,\middle|\, \frac{1}{3}\right) dx \\ &= \int_0^{1/4} \frac{x+\frac{1}{3}}{\frac{1}{3}+\frac{1}{2}} dx = \frac{\frac{1}{32}+\frac{1}{3}}{\frac{1}{3}+\frac{1}{2}} = \frac{14}{32}. \quad \blacksquare \end{aligned}$$

CIMAT

Centro de Investigación en Matemáticas, A.C.

Unidad Monterrey
CIDICS, Campus de la Salud, UANL Ave. Carlos Canseco s/n
Col Mitras Centro, Monterrey N.L., México, C.P.66460 México
Tel +52 (81) 8329 4000 Ext 1778
cimat@cimat.mx / www.cimat.mx