Distribuciones de probabilidad II

Facultad de Estudios Superiores Acatlán

Licenciatura en Economía

Estadística II

20 de abril 2017

José A. Huitrón Mendoza

Distribuciones de probabilidad de Poisson

Enmarca el estudio de una variable aleatoria discreta que se suele usar para estimar el número de veces que sucede un hecho determinado (ocurrencias) en un intervalo de tiempo o de espacio.

Ejemplo: el número de automóviles que llegan a una caseta de cobro.

Si se satisfacen las condiciones (propiedades) siguientes en el número de ocurrencias de una variable discreta, se puede estudiar mediante una **distribución de probabilidad Poisson**.

PROPIEDADES DE UN EXPERIMENTO DE POISSON

1. La probabilidad de ocurrencia es la misma para cualesquiera dos intervalos de la misma magnitud.
2. La ocurrencia o no-ocurrencia en cualquier intervalo es independiente de la ocurrencia o no-ocurrencia en cualquier otro intervalo.

La **función de probabilidad de Poisson** se define mediante la ecuación

FUNCIÓN DE PROBABILIDAD DE POISSON

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

en donde

$f(x) =$ probabilidad de x ocurrencias en un intervalo
$\mu =$ valor esperado o número medio de ocurrencias en un intervalo
$e = 2.71828$

Es preciso observar que el número de ocurrencias *x,* no tiene límite superior. Ésta es una variable aleatoria discreta que toma los valores de una sucesión infinita de números (x = 0,1,2,…).

Vamos a ejemplo de pizarrón…

Aunque esta probabilidad se obtuvo evaluando la función de probabilidad con μ = 10 y x = 5, suele se más sencillo consultar una tabla de probabilidad de Poisson.

Dichas tablas proporcionan las probabilidades para valores específicos de x y μ

Ver ejemplo de tabla

Una *propiedad* de la distribución Poisson es que la media y la varianza son iguales. Por tanto, la varianza del número de llegadas en un lapso de 15 min es $\sigma^2 = 10$, mientras que la desviación estándar es: $\sigma = \sqrt{10} = 3.16.$

Distribución de probabilidad hipergeométrica

Esta distribución está estrechamente relacionada con la distribución binomial.

Pero difieren en dos puntos:

Los ensayos no son independientes

La probabilidad de éxito varía de ensayo a ensayo

Esta función de probabilidad se usa para calcular la probabilidad de que en una muestra aleatoria de n elementos, seleccionados sin reemplazo, se tengan x éxitos y n-x fracasos.

FUNCIÓN DE PROBABILIDAD HIPERGEOMÉTRICA

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \qquad \text{para } 0 \leq x \leq r$$

donde

$f(x)$ = probabilidad de x éxitos en n ensayos

n = número de ensayos

N = número de elementos en la población

r = número de elementos en la población considerados como éxitos

Distribuciones de probabilidad continua

Una diferencia fundamental entre las variables aleatorias discretas y las variables continuas es cómo se calculan las probabilidades.

En las variables aleatorias discretas la función de probabilidad *f(x)* da la probabilidad de que la variable aleatoria tome un valor determinado.

En las variables aleatorias continuas, la contraparte de la función de probabilidad es la función de densidad de probabilidad, que también se denota *f(x).*

La diferencia está en que la función de densidad de probabilidad no da probabilidades directamente.

Es el área bajo la curva f(x) que corresponde a un intervalo determinado proporciona la probabilidad de que la variable aleatoria tome uno de los valores de ese intervalo.

De tal suerte que…

Cuando se calculan probabilidades de variables aleatorias continuas se calcula la probabilidad de que la variable aleatoria tome alguno de los valores dentro de un intervalo.

Distribución de probabilidad uniforme

Consideremos el valor agregado de una empresa en un día de producción, que puede ser cualquier valor en un intervalo que va de 100,000 a 110,000 pesos, de acuerdo con sus capacidades productivas.

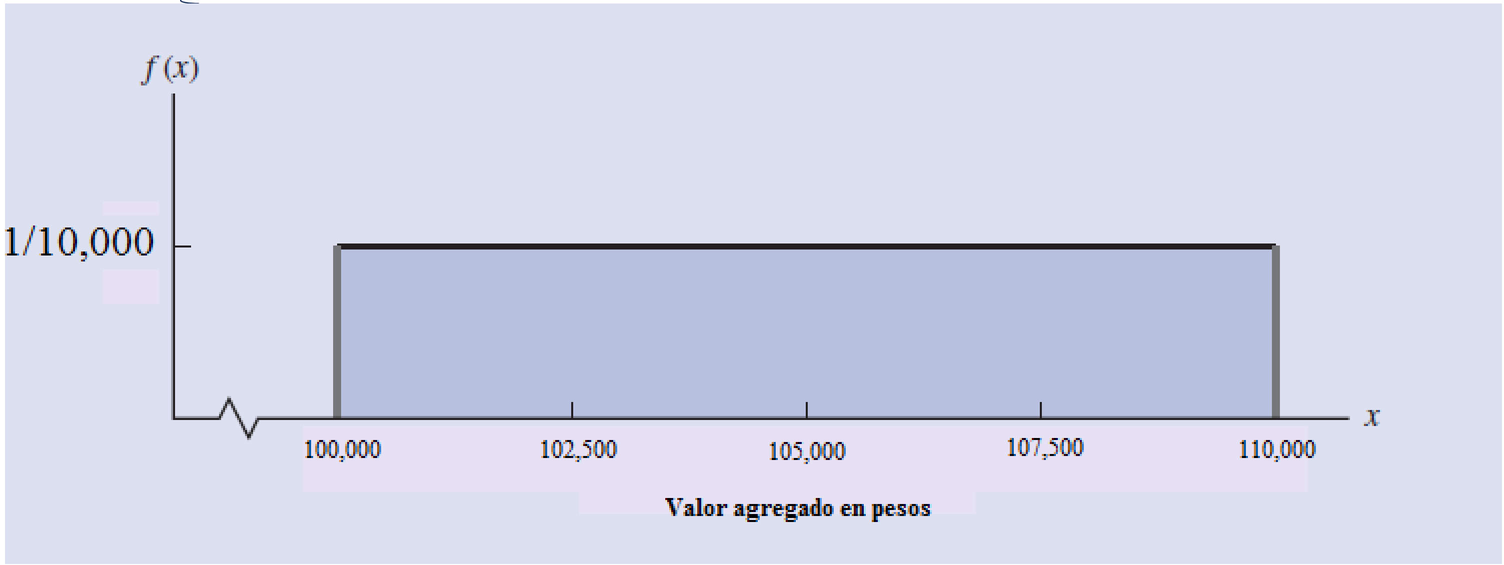
Dado que la variable x toma cualquier valor en ese intervalo, x es una variable continua y no una variable aleatoria discreta.

Contamos con datos suficientes para decir que la probabilidad del VA está en cualquier intervalo de 1 peso.

Como cualquier intervalo de un peso es igual de probable, se dice que la variable aleatoria x tiene una distribución de probabilidad uniforme.

La función de densidad de probabilidad que define la distribución uniforme de la variable aleatoria del valor agregado de la empresa, es:

$$f(x) \begin{cases} = 1/10{,}000 & \text{para } 100{,}000 \leq x \leq 110{,}000 \\ 0 & \text{en cualquier otro caso} \end{cases}$$



¿Cuál es la probabilidad de que el VA se encuentre entre 100,000 y 105,000?

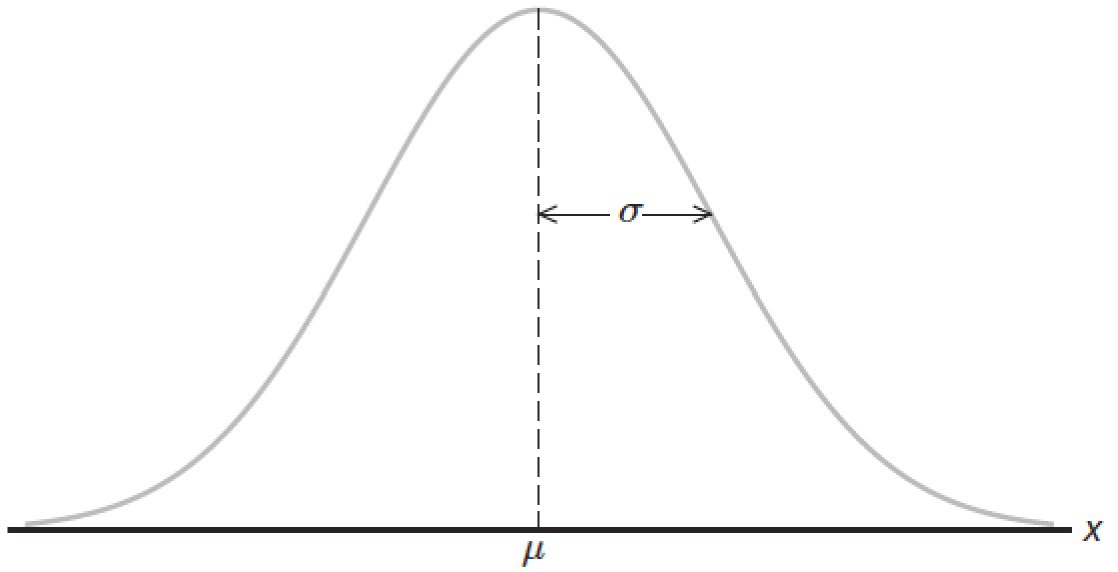
Es factible decir que $P(100{,}000 \leq x \leq 105{,}000) = 0.50$

Distribución normal

La distribución de probabilidad continua más importante en todo el campo de la estadística es la **distribución normal.**

Su gráfica denominada **curva normal,** es una curva con forma de campana, la cual describe de manera aproximada muchos fenómenos que ocurren en la naturaleza, la industria y la investigación.

Una propiedad importante: los errores en las mediciones científicas se aproximan muy bien mediante una distribución normal.



Una variable continua X que tiene la distribución en forma de campana, se denomina variable aleatoria normal.

La ecuación matemática para la distribución de probabilidad de la variable normal depende de los parámetros μ y σ su media y desviación estándar respectivamente. Por ello, denotamos los valores de la densidad de X por $\mathrm{n}(\mathrm{x}; \mu, \sigma)$

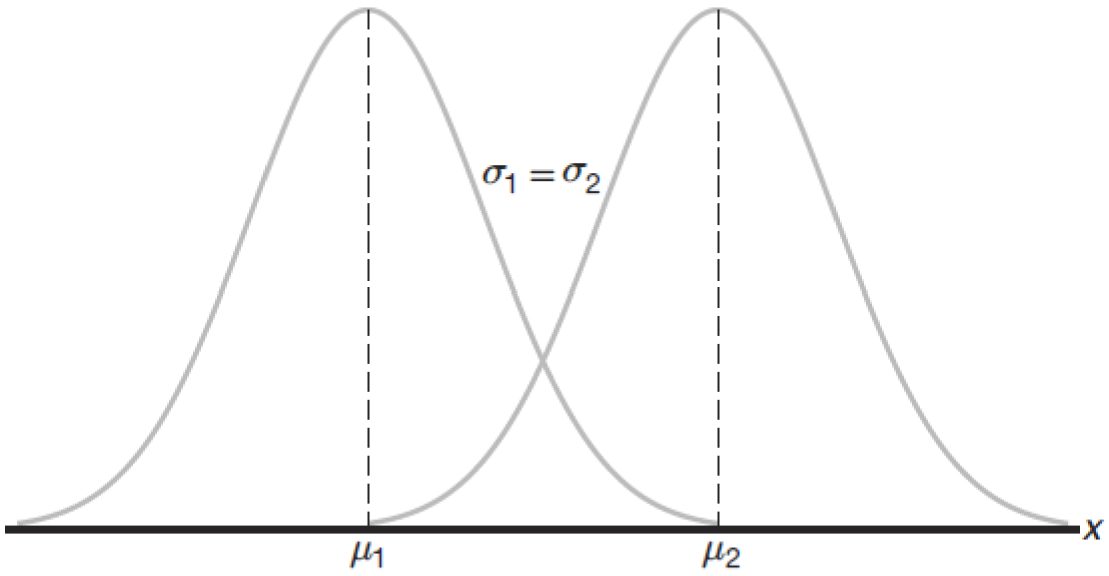
La densidad de la variable aleatoria normal X, con media μ y varianza σ^2, es:

$$n(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \qquad -\infty < x < \infty,$$

donde $\pi = 3.14159\ldots$ y $e = 2.71828\ldots$

Una vez que se especifican μ y σ , la curva queda determinada por completo.

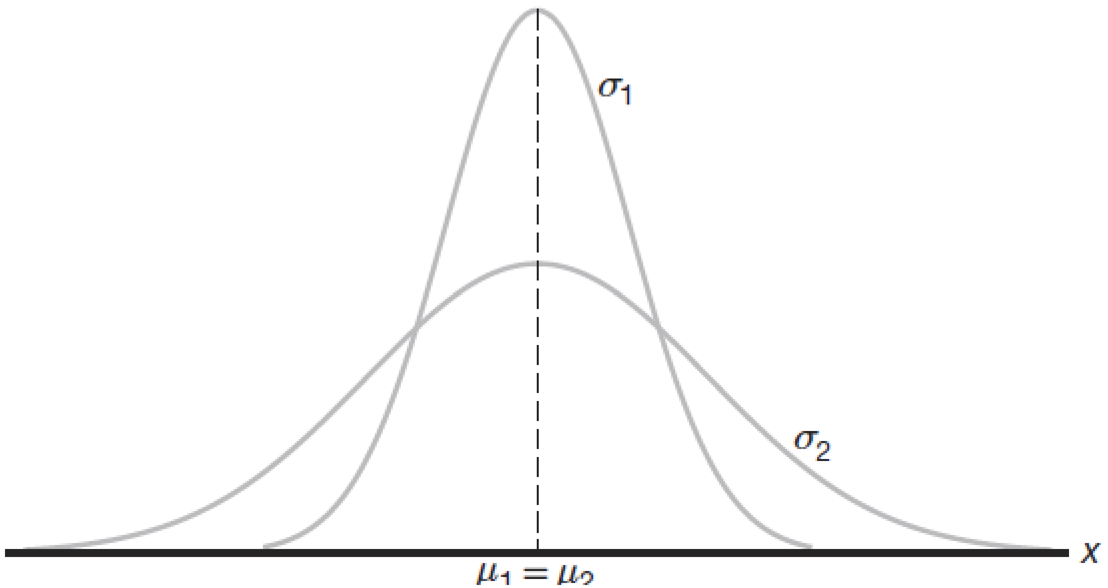
Por ejemplo, si $\mu = 50$ y $\sigma = 5$, entonces se pueden calcular las ordenadas $\mathrm{n}(\mathrm{x}; 50, 5)$, para diferentes valores de x y dibujar la curva.



Curvas normales con $\mu_1 < \mu_2$ y $\sigma_1 = \sigma_2$.

En la figura aparecen dos curvas normales que tienen la misma desviación estándar pero diferentes medias.

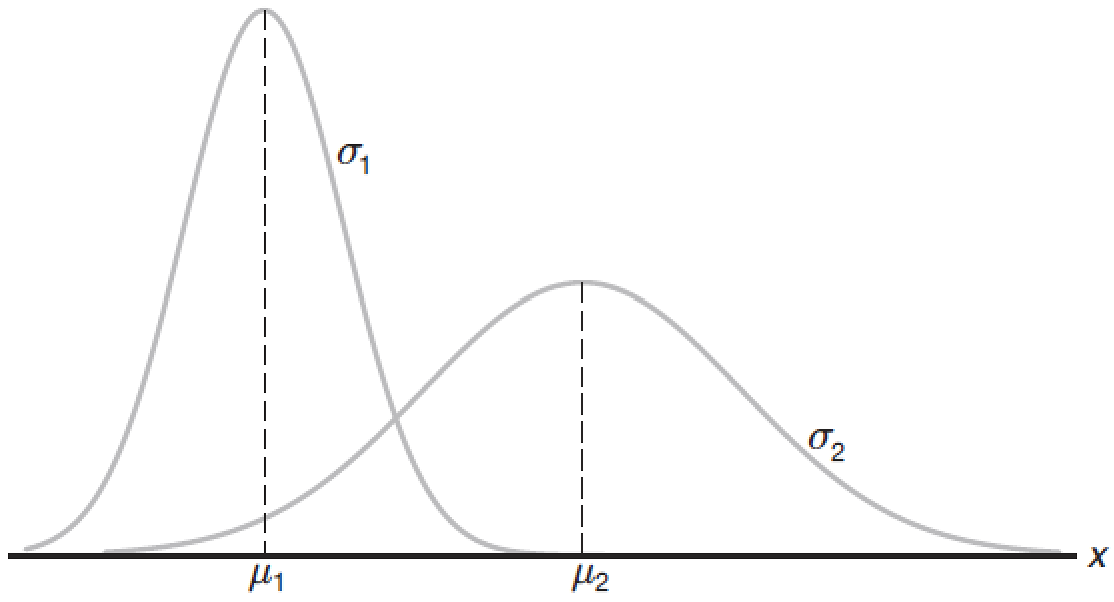
Las dos curvas son idénticas en forma, pero están centradas en diferentes posiciones a lo largo del eje horizontal.



Curvas normales con $\mu_1 = \mu_2$ y $\sigma_1 < \sigma_2$.

En la figura se muestran dos curvas normales con la misma media pero con desviación estándar diferente.

Nota: el área bajo una curva de probabilidad debe ser igual a 1 y, por lo tanto, cuanto más variable sea el conjunto de observaciones, más baja y más ancha será la curva correspondiente.



Curvas normales con $\mu_1 < \mu_2$ y $\sigma_1 < \sigma_2$.

Curvas normales que tienen diferentes medias y diferente desviación estándar.

Con base en lo observado en las gráficas, se listan las siguientes propiedades de la curva normal.

1. La moda, que es el punto sobre el eje horizontal donde la curva tiene su punto máximo, ocurre en $\mathrm{x} = \mu$.
2. La curva es simétrica alrededor de un eje vertical a través de la media μ.
3. La curva tiene sus puntos de inflexión en $\mathrm{x} = \mu \mp \sigma$, es cóncava hacia abajo si $\mu - \sigma < X < \mu + \sigma$, y es cóncava hacia arriba en otro caso.
4. La curva normal se aproxima al eje horizontal de manera asintótica, conforme nos alejamos de la media en cualquier dirección.
5. El área total bajo la curva y sobre el eje horizontal es igual a uno.