Algunas Distribuciones Continuas de Probabilidad

UCR – ECCI

CI-0115 Probabilidad y Estadística

Prof. Kryscia Daviana Ramírez Benavides

Introducción

- El comportamiento de una variable aleatoria queda descrito por su distribución de probabilidad.
  - A menudo, las observaciones de diferentes experimentos estadísticos tienen el mismo tipo general de comportamiento.
  - En consecuencia, las variables aleatorias continuas asociadas se pueden describir con la misma distribución de probabilidad y se pueden representar mediante una sola fórmula.

Distribución Uniforme Continua

- Es la más simple de todas las distribuciones continuas de probabilidad.
- Se caracteriza por una función de densidad que es “plana”, y por ello la probabilidad es uniforme en un intervalo cerrado.
- La función de densidad de la variable aleatoria uniforme continua *X* en el intervalo [*A*,*B*] es

$$f(x;A,B)=\begin{cases}\dfrac{1}{B-A} & A\le x\le B\\ 0 & \text{otro caso}\end{cases}$$

- La media y la varianza de la distribución uniforme continua *f*(*x*;*A*,*B*) son $\mu=\dfrac{A+B}{2}$ $\sigma^2=\dfrac{(B-A)^2}{12}$

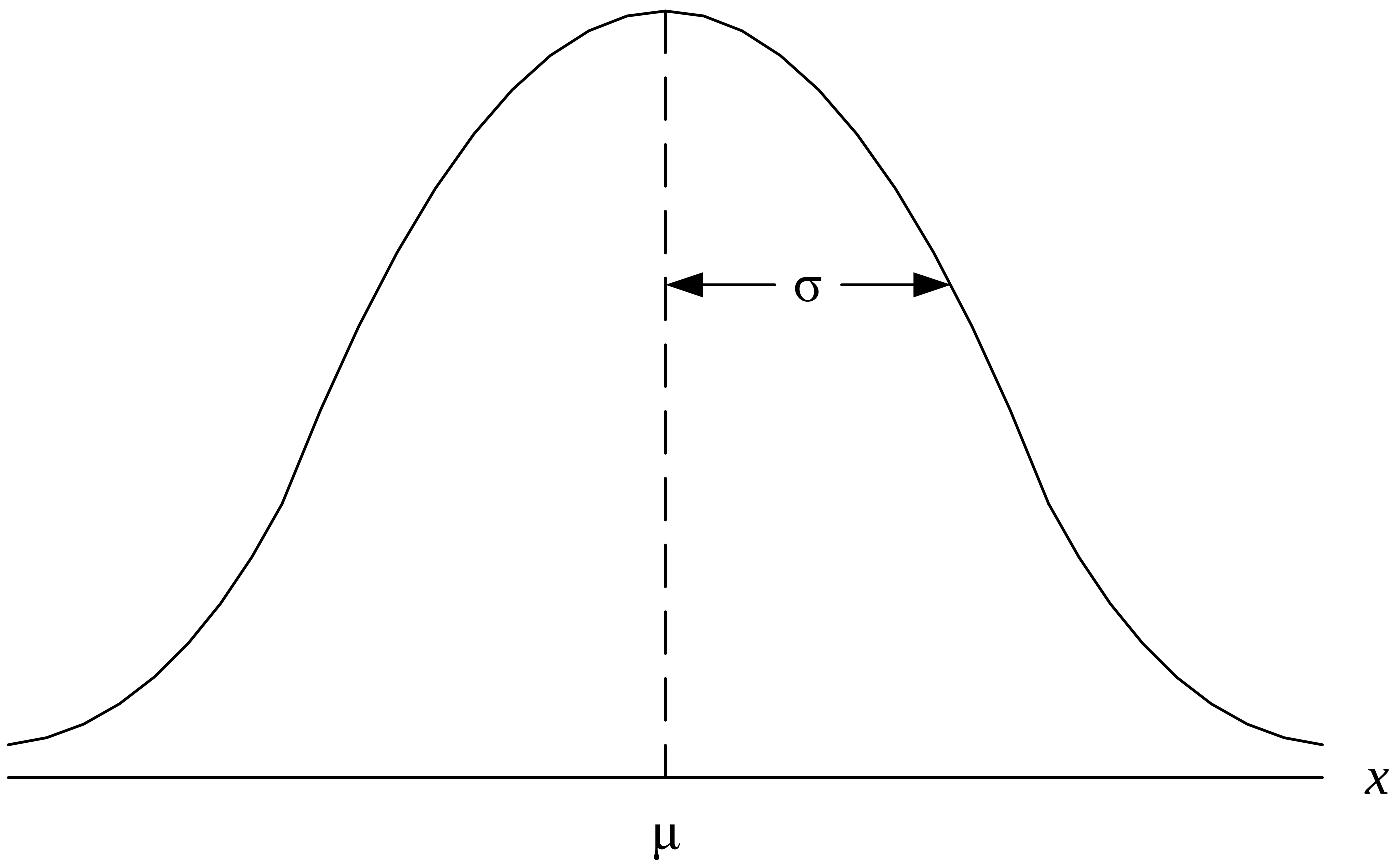
Distribución Uniforme Continua (cont.)

- La función de densidad forma un rectángulo con base $B - A$ y altura constante $1/(B - A)$.
- Como resultado, la distribución uniforme a menudo se llama **distribución rectangular**.
- Es sencillo calcular las probabilidades para la distribución uniforme debido a la naturaleza simple de la función de densidad.
- Sin embargo, la aplicación de esta distribución se basa en la suposición de que la probabilidad de caer en un intervalo de longitud fija dentro de $[A,B]$ es constante.

Distribución Normal

- La distribución continua de probabilidad más importante en todo el campo de la estadística es la **distribución normal**.
- Su gráfica se denomina **curva normal**, es la curva con la forma de campana, la cual describe aproximadamente muchos fenómenos que ocurren en la naturaleza, la industria y la investigación.
  - Las mediciones físicas en áreas como los experimentos meteorológicos, estudios de lluvia y mediciones de partes fabricadas a menudo se explican mejor con una distribución normal.
  - Los errores en las mediciones científicas se aproximan extremadamente bien mediante una distribución normal.

Distribución Normal (cont.)



Distribución Normal (cont.)

- La distribución normal a menudo se denomina **distribución gaussiana**, en honor de Karl Friedrich Gauss.
- Una variable aleatoria continua X que tiene la distribución en forma de campana se llama **variable aleatoria normal**.
- La ecuación matemática para la distribución de probabilidad de la variable normal depende de los dos parámetros μ y σ, su media y desviación estándar, respectivamente.
- De aquí se denota los valores de la densidad de X con $n(x;\mu,\sigma)$.

Distribución Normal (cont.)

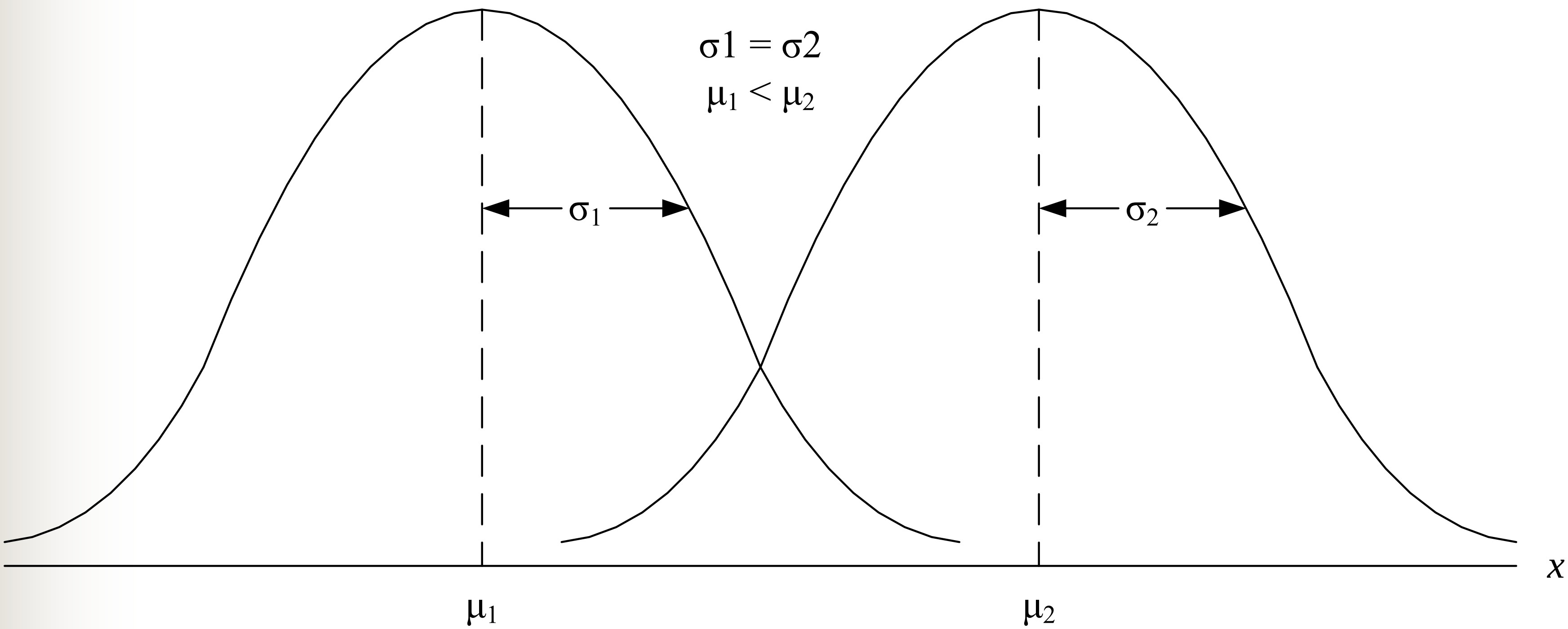
- La función de densidad de la variable aleatoria normal *X*, con media μ y varianza σ^2, es

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

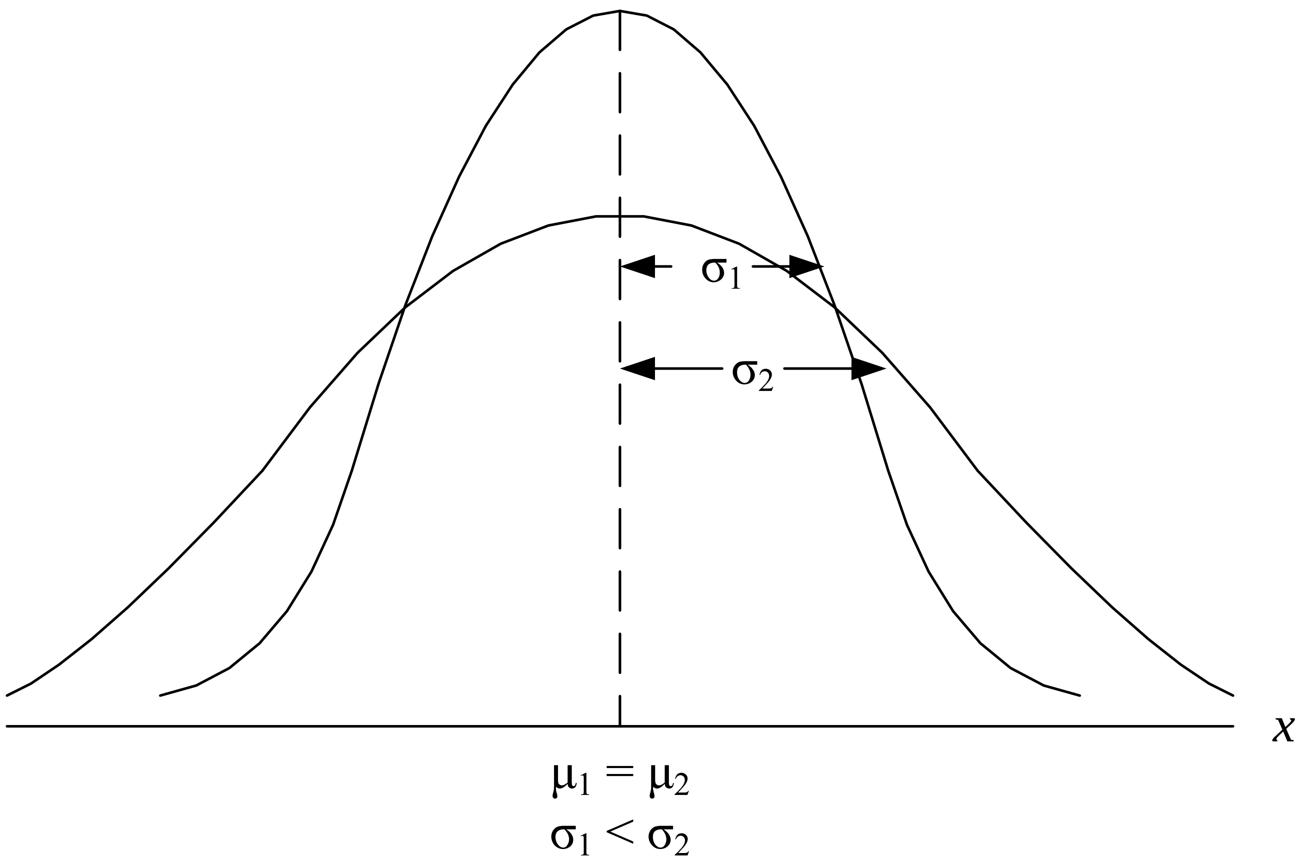
donde $\pi = 3.14159\ldots$ y $e = 2.71828\ldots$

- Una vez que se especifican μ y σ, la curva normal queda determina por completo.

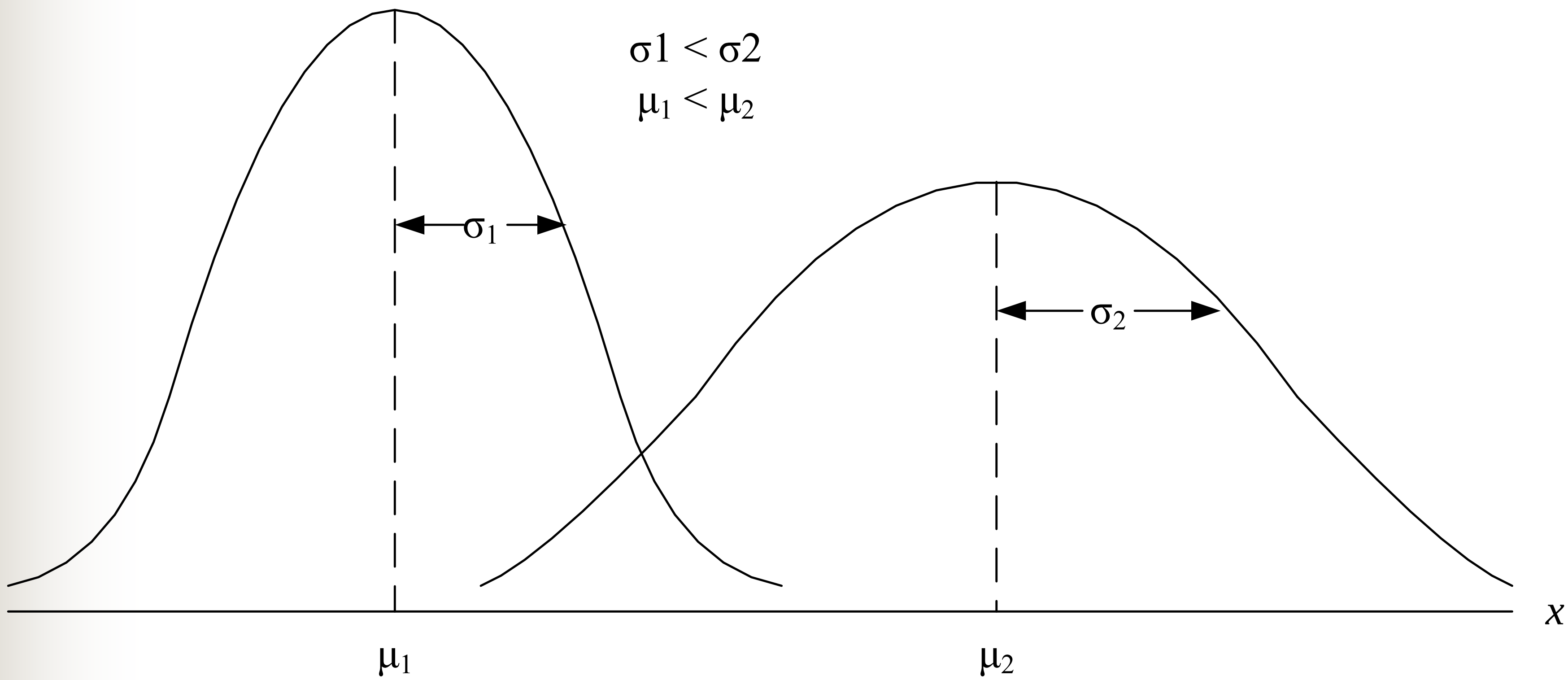
Distribución Normal (cont.)



Distribución Normal (cont.)



Distribución Normal (cont.)



Distribución Normal (cont.)

- Propiedades de la curva normal:
  - La moda, que es el punto sobre el eje horizontal donde la curva es un máximo, ocurre en $x = \mu$.
  - La curva es simétrica alrededor de un eje vertical a través de la media μ.
  - La curva tiene sus puntos de inflexión en $x = \mu \pm \sigma$, es cóncava hacia abajo si $\mu - \sigma < X < \mu + \sigma$, y es cóncava hacia arriba en cualquier otro punto.
  - La curva se aproxima al eje horizontal de manera asintótica conforme se aleja de la media en cualquier dirección.
  - El área total bajo la curva y sobre el eje horizontal es igual a 1.

Distribución Normal (cont.)

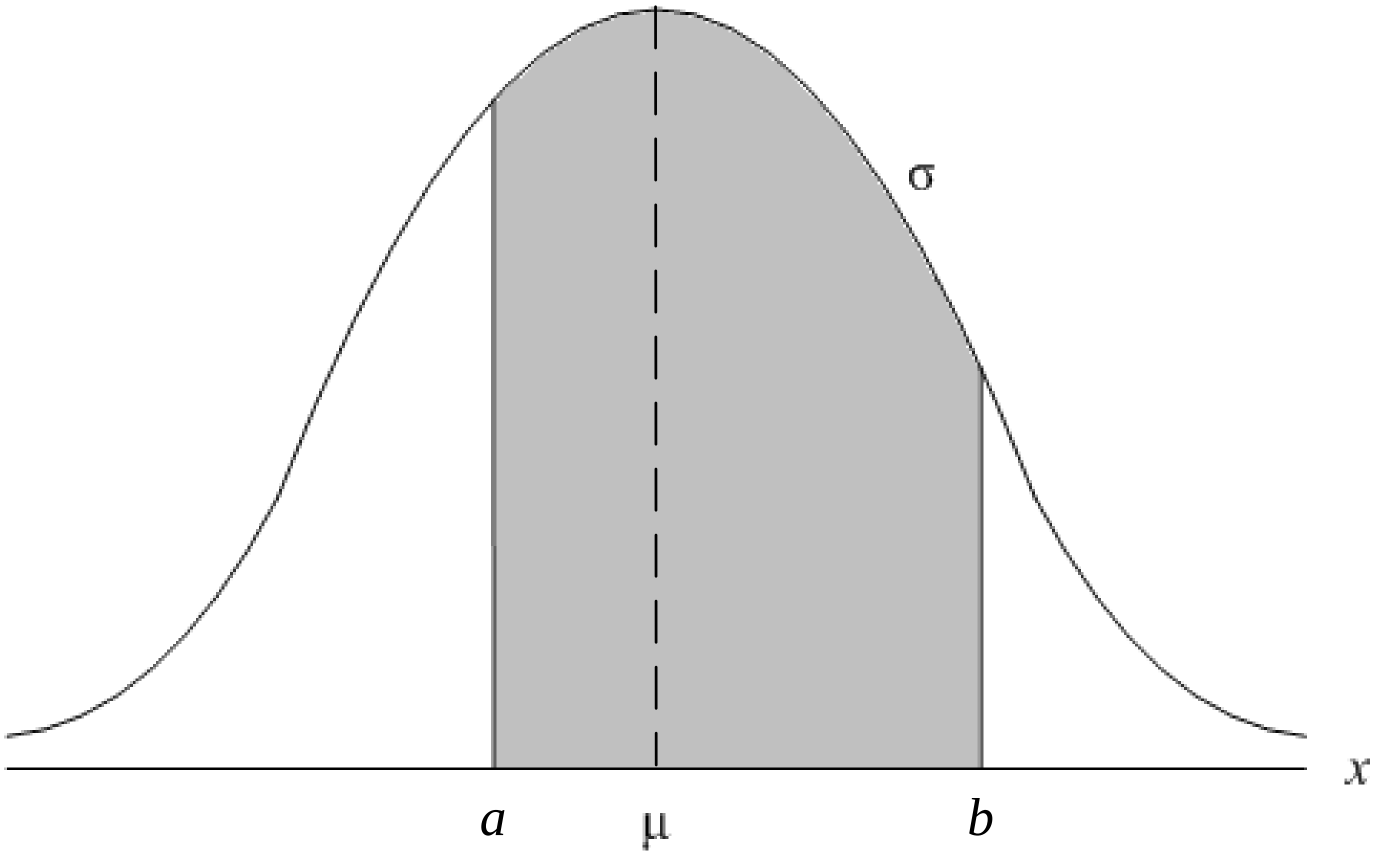
- La distribución normal encuentra una gran aplicación como distribución limitante.
- Bajo ciertas condiciones la distribución normal proporciona una buena aproximación continua a las distribuciones binomial e hipergeométrica.
- La distribución limitante proporciona una base amplia para la inferencia estadística que es muy valiosa para el analista de datos interesado en la estimación y prueba de hipótesis.
- Las áreas importantes del análisis de varianza y del control de calidad tienen su teoría basada en suposiciones que hacen uso de la distribución normal.

Áreas Bajo la Curva Normal

- La curva de cualquier distribución continua de probabilidad o función de densidad se construye de modo que el área bajo la curva limitada por las dos ordenadas $x = a$ y $x = b$ es igual a la probabilidad de que la variable aleatoria X tome un valor entre $x = a$ y $x = b$.

$$P(a < X < b) = \int_a^b n(x;\mu,\sigma)dx = \frac{1}{\sigma\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Áreas Bajo la Curva Normal (cont.)



Áreas Bajo la Curva Normal (cont.)

- El área bajo la curva entre cualesquiera dos valores también depende de los valores μ y σ.
- La dificultad que se encuentra al resolver las integrales de funciones de densidad normal necesita de la tabulación de las áreas de la curva normal para una referencia rápida.
- Sería una tarea sin fin intentar establecer tablas separadas para cada valor concebible de μ y σ.
- Lo bueno es que se puede transformar un conjunto de observaciones de cualquier variable aleatoria normal X a un nuevo conjunto de observaciones de una variable aleatoria normal Z con $\mu = 0$ y $\sigma^2 = 1$.

Áreas Bajo la Curva Normal (cont.)

- Esto se puede realizar por medio de la transformación

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

- Siempre que *X* tome un valor *x*, el valor correspondiente de *Z* está dado por $z = (x - \mu)/\sigma$.
- Por lo tanto, si *X* cae entre los valores *a* y *b*, la variable aleatoria *Z* caerá entre los valores correspondientes $z_1 = (a - \mu)/\sigma$ y $z_2 = (b - \mu)/\sigma$.

Áreas Bajo la Curva Normal (cont.)

- En consecuencia

$$P(a < X < b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} dx$$

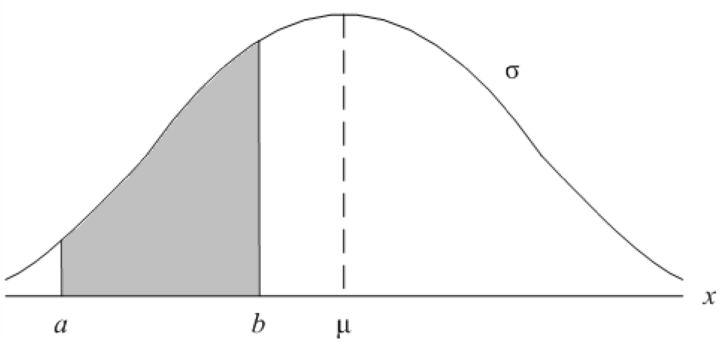
$$= \int_{z_1}^{z_2} n(z;0,1) dx$$

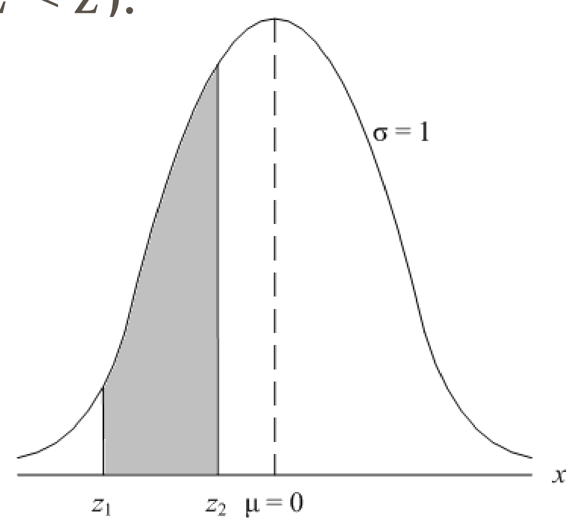
$$P(a < X < b) = P(z_1 < Z < z_2)$$

donde *Z* se ve como una variable aleatoria normal con $\mu = 0$ y $\sigma^2 = 1$.

Áreas Bajo la Curva Normal (cont.)

- La distribución de una variable aleatoria normal con $\mu = 0$ y $\sigma^2 = 1$ se llama **distribución normal estándar**.
- Ver Tabla A.3 del libro de texto para los valores calculados del área bajo la curva normal, $P(Z < z)$.





Áreas Bajo la Curva Normal (cont.)

- De acuerdo con el teorema de Chebyshev, la probabilidad de que una variable aleatoria tome un valor dentro de dos desviaciones estándar de la media es al menos ¾.
- Si la variable aleatoria tiene una distribución normal, los valores *z* que corresponde a $a = \mu - 2\sigma$ y $b = \mu + 2\sigma$ se calculan fácilmente y son

$$z_1 = \frac{(\mu - 2\sigma) - \mu}{\sigma} = -2 \qquad z_2 = \frac{(\mu + 2\sigma) - \mu}{\sigma} = 2$$

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = P(-2 < Z < 2) = P(Z < 2) - P(Z < -2)$$

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9772 - 0.0228 = 0.9544$$

Áreas Bajo la Curva Normal (cont.)

- Si se desea saber los valores de la variable aleatoria *X*, a partir de los valores de la variable aleatoria *Z*, se utiliza la fórmula

$$x = z\sigma + \mu$$

Aproximación Normal a la Binomial

- La distribución normal a menudo es una buena aproximación a una distribución discreta cuando la última adquiere una forma de campana simétrica.
- Desde el punto de vista lógico, algunas distribuciones convergen a la normal conforme sus parámetros se aproximan a ciertos límites.
- La distribución normal no solo proporciona una aproximación muy precisa para valores grandes de n y p, no está extremadamente cercana de 0 o 1 (como Poisson), sino que también proporciona una aproximación bastante buena aun cuando n es pequeña y p está razonablemente cercana a ½.

Aproximación Normal a la Binomial (cont.)

- Si X es una variable aleatoria binomial con $\mu = np$ y $\sigma^2 = npq$, entonces la forma limitante de la distribución de

$$Z = \frac{X - np}{\sqrt{npq}}$$

conforme $n \to \infty$, es la distribución normal estándar $n(z;0,1)$.

Distribuciones Gamma y Exponencial

- Las distribuciones gamma y exponencial juegan un papel importante en teoría de colas y problemas de confiabilidad.
- Los tiempos entre llegadas en instalaciones de servicio, y tiempo de falla de componentes y sistemas eléctricos, a menudo quedan bien modeladas con la distribución exponencial.
- La distribución exponencial es un caso especial de la distribución gamma.

Distribuciones Gamma y Exponencial (cont.)

- La distribución gamma deriva su nombre de la bien conocida **función gamma**, que se estudia en muchas áreas de las matemáticas.
- La **función gamma** se define como

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx \quad a > 0$$

$$\Gamma(1) = 1 \quad \Gamma(n) = (n-1)!$$

- Una propiedad importante de la función gamma es que

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Distribuciones Gamma y Exponencial (cont.)

- La variable aleatoria continua *X* tiene una **distribución gamma**, con parámetros α y β, si su función de densidad está dada por

$$f(x) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & \text{otro caso} \end{cases}$$

  donde $\alpha > 0$ y $\beta > 0$.

- La media y la varianza de la distribución gamma son

$$\mu = \alpha\beta \qquad \sigma^2 = \alpha\beta^2$$

Distribuciones Gamma y Exponencial (cont.)

- Cuando la distribución gamma tiene $\alpha = 1$ se llama **distribución exponencial**.
- La variable aleatoria continua *X* tiene una **distribución exponencial**, con parámetro β, si su función de densidad está dada por

$$f(x) = \begin{cases} \dfrac{1}{\beta} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & \text{otro caso} \end{cases}$$

  donde $\beta > 0$.
- La media y la varianza de la distribución gamma son

$$\mu = \beta \quad \sigma^2 = \beta^2$$

Distribuciones Gamma y Exponencial (cont.)

- Las aplicaciones más importantes de la distribución exponencial son en situaciones donde se aplica el proceso de Poisson.
- La distribución de Poisson se utiliza para calcular la probabilidad de números específicos de eventos durante un **periodo** o **espacio** particular. En muchas aplicaciones, el tiempo o la cantidad de espacio es la variable aleatoria.
- La distribución exponencial se utiliza como modelo para representar tiempo de funcionamiento o de espera. También expresa el tiempo que transcurre entre sucesos que se contabilizan mediante la distribución de Poisson.

Distribuciones Gamma y Exponencial (cont.)

- La relación entre la distribución exponencial y el proceso de Poisson es simple.
- El parámetro λ de Poisson se puede interpretar como el número medio de eventos por unidad de "tiempo".
- Considerar la variable aleatoria descrita por el tiempo que se requiere para que ocurra el primer evento.
- Con el uso de Poisson, se encuentra que la probabilidad de que no ocurra algún evento, en el periodo hasta el tiempo *t* está dada por

$$p(0;\lambda t)=\frac{e^{-\lambda t}(\lambda t)^0}{0!}=e^{-\lambda t}$$

Distribuciones Gamma y Exponencial (cont.)

- Se puede utilizar lo anterior y hacer que X sea el tiempo para el primer evento de Poisson.
- La probabilidad de que la duración del tiempo hasta el primer evento exceda x es la misma que la probabilidad de que no ocurra algún evento de Poisson en x.
- Como resultado $P(X \geq x) = e^{-\lambda t}$
- Así, la función de distribución acumulada para X está dada por

$$F(x) = P(0 \leq X \leq x) = 1 - e^{-\lambda t}$$

Distribuciones Gamma y Exponencial (cont.)

- Ahora, a fin de reconocer la presencia de la distribución exponencial se puede diferenciar la función de distribución acumulada anterior para obtener la función de densidad

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

  que es la función de densidad de la distribución exponencial con $\lambda = 1/\beta$.

Distribuciones Gamma y Exponencial (cont.)

- Si *X* v.a. con distribución exponencial parámetro λ, entonces:
  - La función de distribución de densidad de *X* es:

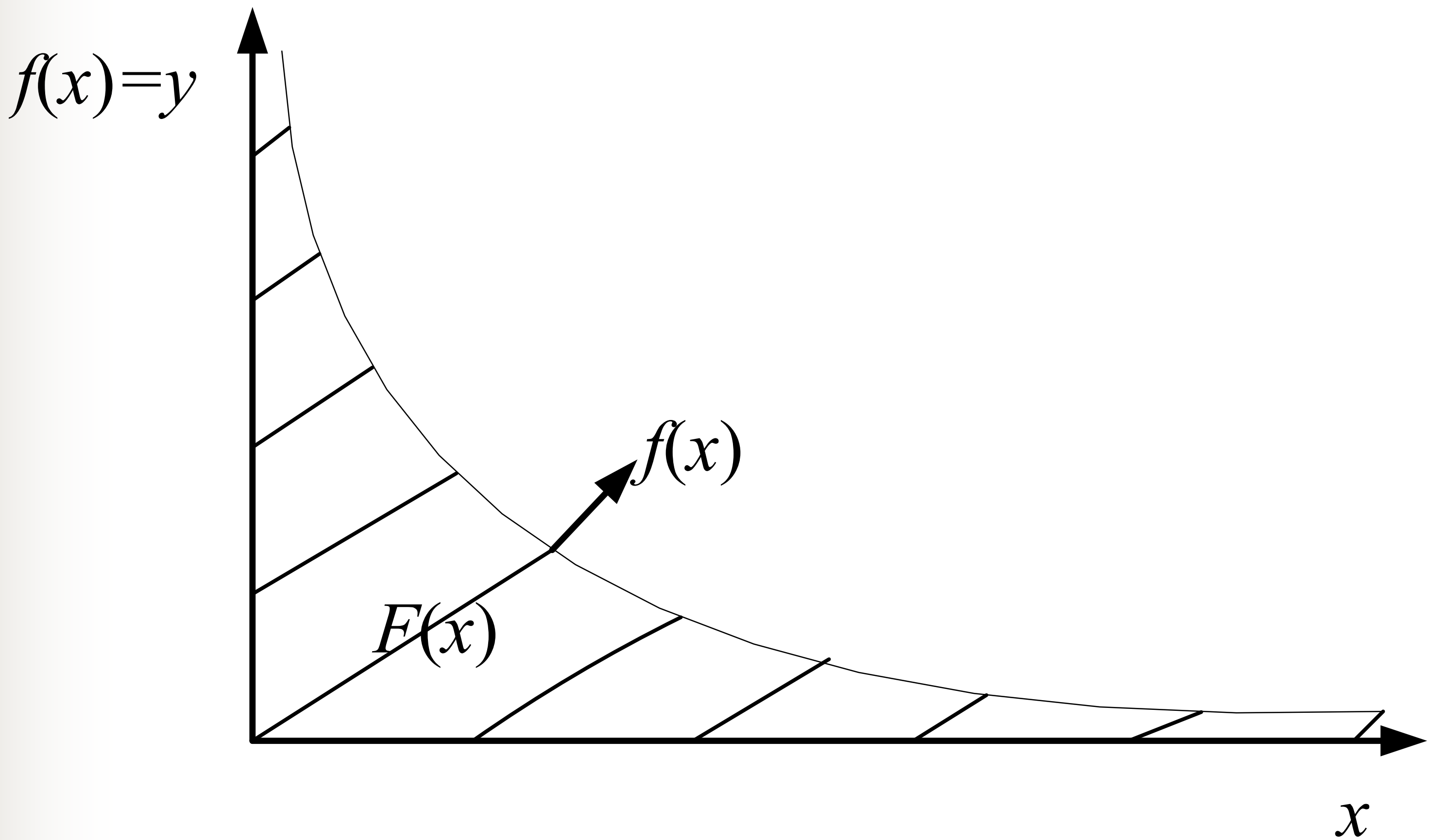
$$f(x) = \lambda e^{-\lambda x} = \frac{e^{-\frac{x}{\beta}}}{\beta}$$

  - La función de distribución acumulada de *X* es:

$$F(x) = 1 - e^{-\lambda x} = 1 - e^{-\frac{x}{\beta}}$$

$$F(x) = P(X \leq x)$$

Distribuciones Gamma y Exponencial (cont.)



Distribuciones Gamma y Exponencial (cont.)

- El parámetro β es el tiempo medio entre eventos. En teoría de confiabilidad, donde la falla del equipo a menudo se ajusta al proceso de Poisson, este parámetro se llama **tiempo medio entre fallas**.
- La distribución gamma también se aplica en aplicaciones importantes en tiempo de espera y teoría de confiabilidad.
  - Mientras que la **distribución exponencial** describe el **tiempo (o espacio) que transcurre hasta la ocurrencia de un evento de Poisson (o el tiempo entre eventos)**, la **distribución gamma** describe el **tiempo (o espacio) que transcurre hasta que ocurre un número específico de eventos de Poisson**.

Distribuciones Gamma y Exponencial (cont.)

- Mientras que la distribución gamma trata con el tiempo (o espacio) hasta la ocurrencia de α eventos de Poisson, hay otros usos donde esta distribución trabaja muy bien aunque no exista una estructura de Poisson clara.
- Uno de estos usos es para problemas de **tiempo de sobrevivencia** en aplicaciones de ingeniería y biomédicas.

Distribuciones Gamma y Exponencial (cont.)

- La **función gamma incompleta** resulta ser la función de distribución acumulada de la gamma.

$$F(x;\alpha) = \int_0^x \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)}dy$$

$$y = \frac{X}{\beta} \quad X = \beta y$$

- En la tabla A.24 está la función gamma incompleta.

Distribución ji Cuadrada – Chi Cuadrada

- Otro caso especial importante de la distribución gamma se obtiene al hacer $\alpha = v/2$ y $\beta = 2$, donde v es un entero positivo.
- Este caso especial es la **distribución ji cuadrada**. Este distribución tiene un solo parámetro, v, llamado grados de libertad.
- Esta distribución juega un papel muy importante en la inferencia estadística; tiene una aplicación considerable en la metodología y en la teoría.
- Además, es un componente importante de la prueba de hipótesis y la estimación estadística.

Distribución ji Cuadrada – Chi Cuadrada (cont.)

- La variable aleatoria continua *X* tiene una **distribución ji cuadrada**, con *v* grados de libertad, si su función de densidad está dada por

$$f(x)=\begin{cases}\dfrac{1}{2^{\frac{v}{2}}\Gamma\left(\dfrac{v}{2}\right)}x^{\frac{v}{2}-1}e^{-\frac{x}{2}} & x>0\\ 0 & \text{otro caso}\end{cases}$$

  donde *v* es un entero positivo.

- La media y la varianza de la distribución ji cuadrada son

$$\mu=v \quad \sigma^2=2v$$

Distribución Logarítmica Normal

- La distribución logarítmica normal se utiliza en una amplia variedad de aplicaciones.
- Se aplica en casos donde una transformación de logarítmico natural tiene como resultado una distribución normal.

Distribución Logarítmica Normal (cont.)

- La variable aleatoria continua *X* tiene una **distribución logarítmica normal** si la variable aleatoria *Y* = ln(*X*) tiene una distribución normal con media μ y desviación estándar σ. La función de densidad de *X* que resulta es

$$f(x) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}} & x \geq 0 \\ 0 & \text{otro caso} \end{cases}$$

- La media y la varianza de la distribución logarítmica normal son

$$\mu = e^{\mu + \frac{\sigma^2}{2}} \qquad \sigma^2 = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$$

Distribución de Weibull

- Cuando un componente está en un periodo de investigación, la razón de falla puede no ser constante durante dicho tiempo, lo que nos podría hacer pensar que el periodo de vida inicial del componente fue usado sólo para la investigación.
  - Es entonces donde el tiempo de falla se determina por desgaste y no de manera aleatoria como se necesita.
  - Es aquí donde el modelo exponencial no es aplicable y se hace necesario utilizar una razón de falla más general.
- El modelo de Weibull sería entonces la distribución con una razón de falla más general aplicable en estas circunstancias, ya que describe los tiempos de falla de componentes cuando sus razones de falla crecen o decrecen con el tiempo.

Distribución de Weibull (cont.)

- La variable aleatoria continua *X* tiene una **distribución de Weibull**, con parámetros α y β, si su función de densidad está dada por

$$f(x)=\begin{cases}\alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}} & x>0\\ 0 & \text{otro caso}\end{cases}$$

  donde $\alpha > 0$ y $\beta > 0$.

- La media y la varianza de la distribución de Weibull son

$$\mu=\alpha^{-\frac{1}{\beta}}\Gamma\left(1+\frac{1}{\beta}\right)\qquad \sigma^2=\alpha^{-\frac{2}{\beta}}\left\{\Gamma\left(1+\frac{2}{\beta}\right)-\left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2\right\}$$

Distribución de Weibull (cont.)

- Como la distribución gamma y exponencial, la distribución de Weibull también se aplica a problemas de confiabilidad y de prueba de vida como los **tiempos de falla** o **duración de vida** de un componente, medido en algún tiempo específico hasta que falla.
- El tiempo de falla se representa mediante la variable aleatoria continua T, con función de densidad de probabilidad $f(t)$, que es una distribución de Weibull.

Distribución de Weibull (cont.)

- **Función de Confiabilidad de Weibull.** Se infiere que la función de confiabilidad asociada con la distribución de tiempos de falla de Weibull es:

$$R(t)=P(T>t)=\int_t^{+\infty} f(t)dt=1-F(t)=e^{-\alpha t^{\beta}} \qquad t>0$$

- **Función Razón de Falla de Weibull.** La razón de falla que conduce a la distribución de Weibull esta dada por:

$$Z(t)=\frac{F(t)'}{R(t)}=\frac{f(t)}{1-F(t)}=\frac{F(t)'}{1-F(t)}=\frac{f(t)}{R(t)}=\alpha\beta t^{\beta-1} \qquad t>0$$

Distribución de Weibull (cont.)

- La tasa de falla disminuye con el tiempo si $\beta < 1$, aumenta con el tiempo si $\beta > 1$, y es constante si $\beta = 1$.
- La distribución exponencial se aplica para una tasa de falla constante (vía el proceso de Poisson).

Referencias Bibliográficas

- Walpole, R.E.; Myers, R.H.; Myers, S.L. & Ye, K. "Probabilidad y estadística para ingeniería y ciencias". Octava Edición. Pearson Prentice-Hall. México, 2007.