



Distribuciones de probabilidad de variables continuas

PARAGUAYO ALEMANA

# Distribuciones de probabilidad de variables continuas



#### Libro de texto

 Probability and Statistics for Engineering and the Sciences, Ninth Edition, Jay L. Devore, Cengage Learning. (2014)

### DISTRIBUCIÓN DE PROBABILIDAD DE VARIABLE CONTINUAS

PARAGUAYO ALEMANA

#### Objetivos de aprendizaje:

Al terminar este capítulo, ustedes serán capaces de:

- 1. Conocer la distribución normal.
- 2. Identificar las propiedades de una distribución normal.
- 3. Calcular probabilidades en distribuciones normales.
- 4. Encontrar el área bajo una distribución normal estándar.
- 5. Interpretar áreas bajo la curva normal de acuerdo al problema.



- Una variable aleatoria es un valor numérico que corresponde al resultado de un experimento aleatorio, como por ejemplo:
  - el número de caras que se obtienen al lanzar 3 veces una moneda
  - el número de lanzamientos de un dado hasta que aparece el seis
  - el número de llamadas que se reciben en un teléfono en una hora
  - el tiempo de espera a que llegue un autobús
- Las variables aleatorias pueden ser discretas
   continuas.



#### VARIABLE ALEATORIA

- EJEMPLO 1: considere un experimento aleatorio en el que se lanza tres veces una moneda. Sea X el número de caras. Sea H el resultado de obtener una cara y T el de obtener una cruz.
- El *espacio muestral* para este experimento será: *TTT, TTH, THT, THH, HTT, HTH, HHT, HHH.*
- Entonces, los valores posibles de X (número de caras) son x = 0, 1, 2, 3.

- El resultado "cero caras" ocurrió una vez.
- El resultado "una cara" ocurrió tres veces.
- El resultado "dos caras" ocurrió tres veces.
- El resultado "tres caras" ocurrió una vez.
- De la definición de variable aleatoria, la *X* definida en este experimento, es una *variable aleatoria*.



#### DISTRIBUCIONES PROBABILISTICAS

- Una distribución probabilística es la enumeración de todos los resultados de un experimento junto con las probabilidades asociadas.
- Para el **EJEMPLO 1**:

Número de Caras	Probabilidad de los Resultados
0	1/8 = 0,125
1	3/8 = 0,375
2	3/8 = 0,375
3	1/8 = 0,125
Total	8/8 = 1



### CARACTERÍSTICAS DE UNA DISTRIBUCIÓN PROBABILÍSTICA

- La probabilidad de un resultado siempre debe estar entre 0 y 1.
- La suma de todos los resultados mutuamente excluyentes siempre es 1.



#### VARIABLE ALEATORIA DISCRETA

- Una variable aleatoria discreta es una variable que puede tomar sólo ciertos valores diferentes que son el resultado de la cuenta de alguna característica de interés.
- **EJEMPLO 2:** sea *X* el número de caras obtenidas al lanzar 3 veces una moneda.
- Aquí los valores de X son x = 0, 1, 2, 3.



#### VARIABLE ALEATORIA CONTINUA

- Una variable aleatoria continua es una variable que puede tomar un número infinito de valores.
- **Ejemplos:** la altura de un jugador de básquetbol o el tiempo que dura un trayecto en autobús.



## MEDIA DE UNA DISTRIBUCIÓN DE PROBABILÍSTICA DISCRETA

PARAGUAYO ALEMANA

#### • La **media**:

- indica la ubicación central de los datos.
- Es el promedio del valor de la variable aleatoria.
- también se conoce como el valor esperado, *E(x),* en una distribución de probabilidad.
- se puede considerar como un promedio ponderado.



## MEDIA DE UNA DISTRIBUCIÓN DE PROBABILÍSTICA DISCRETA

PARAGUAYO ALEMANA

La media se calcula con la fórmula:

$$\mu = E(x) = [x * P(x)]$$

• donde  $\mu$  representa la media y P(x) es la probabilidad de los diferentes resultados x.



## VARIANZA DE UNA DISTRIBUCIÓN DE PROBABILÍSTICA DISCRETA

- La varianza mide la cantidad de dispersión (variación) de una distribución.
- La varianza de una distribución discreta se denota por la letra griega σ² (sigma cuadrada).
- La desviación estándar (σ) se obtiene tomando la raíz cuadrada de σ².



## VARIANZA DE UNA DISTRIBUCIÓN DE PROBABILÍSTICA DISCRETA

PARAGUAYO ALEMANA

 La varianza de una distribución de probabilidad discreta se calcula a partir de la fórmula

$$\sigma^2 = \sum [(x - \mu)^2 * P(x)]$$



 Dan Desch, propietario de College Painters, estudió sus registros de las últimas 20 semanas y obtuvo los siguientes números de casas pintadas por semana:

# Casas Pintadas	Semanas
10	5
11	6
12	7
13	2



Distribución probabilística:

Número de Casas Pintadas, <i>X</i>	Probabilidad <i>P(X)</i>
10	0,25
11	0,30
12	0,35
13	0,10
Total	1,0



 Calcule el número medio de casas pintadas por semana:

$$\mu = E(x) = \Sigma[xP(x)]$$

$$= (10)(.25) + (11)(.30) + (12)(.35) + (13)(.10)$$

$$= 11.3$$



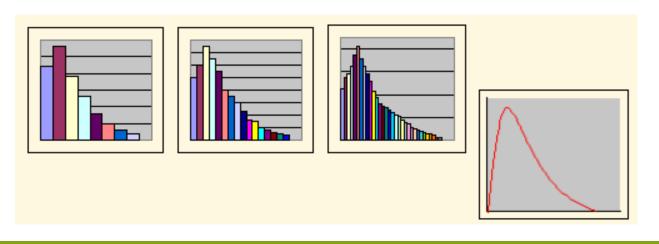
 Calcule la varianza del número de casas pintadas por semana:

$$\sigma^{2} = \Sigma[(x - \mu)^{2} P(x)]$$
=.4225+.0270+.1715+.2890
=.91



## DISTRIBUCIONES DE PROBABILIDAD CONTINUAS

- Las distribuciones de probabilidad son idealizaciones de polígonos de frecuencias.
- En el caso de una variable estadística continua consideramos un histograma de frecuencias relativas, y se comprueba que al aumentar el número de datos y el número de clases el histograma tiende a estabilizarse llegando a convertirse su perfil en la gráfica de una función.





- Las distribuciones de probabilidad de variable continua se definen mediante una función y = f(x) llamada función de probabilidad.
- Así como en el histograma la frecuencia viene dada por el área, en la función de probabilidad, la probabilidad viene dada por el área bajo la curva, por lo que:
  - El área encerrada bajo la totalidad de la curva es 1.
  - Para obtener la probabilidad p(a ≤ X ≤ b) obtenemos la proporción de área que hay bajo la curva desde a hasta b.
  - La probabilidad de sucesos puntuales es 0, p(X=a) = 0.



## FUNCIÓN DE DENSIDAD Y FUNCIÓN DE DISTRIBUCIÓN

- Llamaremos función de probabilidad de una variable aleatoria continua X a una función f que cumple:
  - es positiva
  - el área total bajo la curva, es decir entre f(x) y la distancia horizontal al eje vertical, es 1
  - el área determinada por f(x), la distancia horizontal al eje vertical y las rectas x=a, x=b, es la probabilidad de que la variable continua X esté en el intervalo [a,b], p(a≤X≤b)



### PARÁMETROS DE UNA DISTRIBUCIÓN DE PROBABILIDAD

PARAGUAYO ALEMANA

• Por analogía con las variables discretas, podemos definir también aquí la media  $\mu$  y la desviación típica  $\sigma$  de la variable continua.



### PARÁMETROS DE UNA DISTRIBUCIÓN DE PROBABILIDAD

- La media  $\mu$ , es un valor representativo de todos los valores que toma la variable aleatoria X.
- Para calcularla debemos de resolver:

$$\mu = \int_{-\infty}^{+\infty} x f(x) dx$$



### PARÁMETROS DE UNA DISTRIBUCIÓN DE PROBABILIDAD

- La desviación típica  $\sigma$  es una medida de la dispersión de los valores que toma la variable aleatoria respecto de la media.
- La desviación típica será más pequeña o más grande según la gráfica de la función de probabilidad sea más estrecha o más ancha en torno a la media.
- En este caso se calcula:

$$\sigma = \sqrt{\int_{-\infty}^{+\infty} x^2 f(x) dx - \mu^2}$$



- Entre las distribuciones continuas, la más importante es la distribución normal.
- Fue introducida por Carl Friedrich Gauss a principios del siglo XIX en su estudio de los errores de medida.
- Desde entonces se ha utilizado como modelo en multitud de variables (peso, altura, calificaciones...), en cuya distribución los valores más usuales se agrupan en torno a uno central y los valores extremos son escasos.



- Se utiliza a menudo porque hay muchas variables asociadas a fenómenos naturales que siguen el modelo de la distribución normal.
  - Caracteres morfológicos de individuos (personas, animales, plantas,...) de una especie, por ejemplo: tallas, pesos, diámetros, distancias, perímetros,...
  - Caracteres fisiológicos, por ejemplo: efecto de una misma dosis de un fármaco, o de una misma cantidad de abono.



- Caracteres sociológicos, por ejemplo: consumo de cierto producto por un mismo grupo de individuos, puntuaciones de examen.
- Caracteres psicológicos, por ejemplo: cociente intelectual, grado de adaptación a un medio.
- Errores cometidos al medir ciertas magnitudes.
- Valores estadísticos muestrales como la media, varianza y moda.



#### LA FUNCION DE DISTRIBUCION

- Puede tomar cualquier valor (- ∞, + ∞).
- Hay más probabilidad para los valores cercanos a la media μ.
- Conforme nos separamos de μ, la probabilidad va decreciendo de igual forma a derecha e izquierda (es simétrica), dependiendo de la desviación típica σ.



## LA FUNCIÓN F(X)

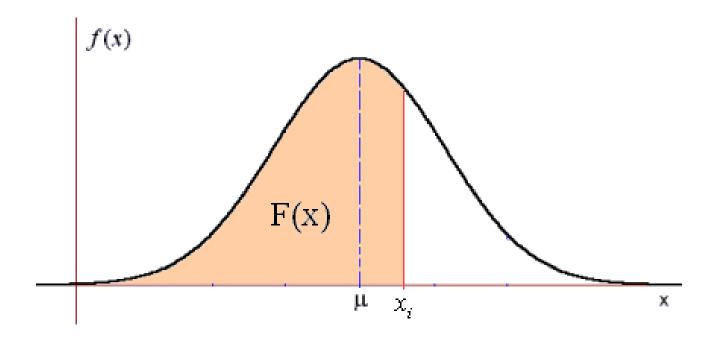
 Una variable aleatoria continua sigue una distribución normal si su función de probabilidad es:

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

- donde  $\mu$  y  $\sigma$  coinciden respectivamente con la **media** y la **desviación típica** de la variable aleatoria.
- Estos parámetros son los que determinan esta distribución que designaremos por  $N(\mu, \sigma)$ .

## LA FUNCIÓN F(X)

• F(x) es el área sombreada de la siguiente gráfica:





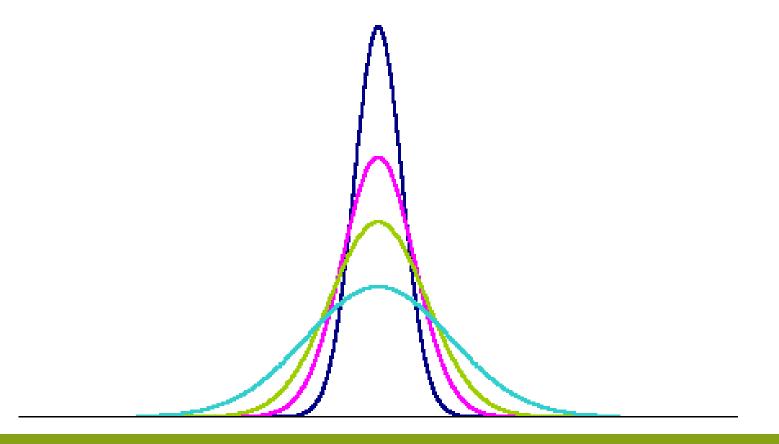
#### PROPIEDADES DE LA DISTRIBUCIÓN NORMAL

- La forma de la campana de Gauss depende de los parámetros  $\mu$  y  $\sigma$ .
- Tiene una única moda que coincide con su media y su mediana.
- La curva normal es asintótica al eje de X.
- Es simétrica con respecto a su media  $\mu$ .
- Según esto, para este tipo de variables existe una probabilidad de un 50% de observar un dato mayor que la media, y un 50% de observar un dato menor.

### LA DESVIACIÓN ESTÁNDAR $(\sigma)$

PARAGUAYO ALEMANA

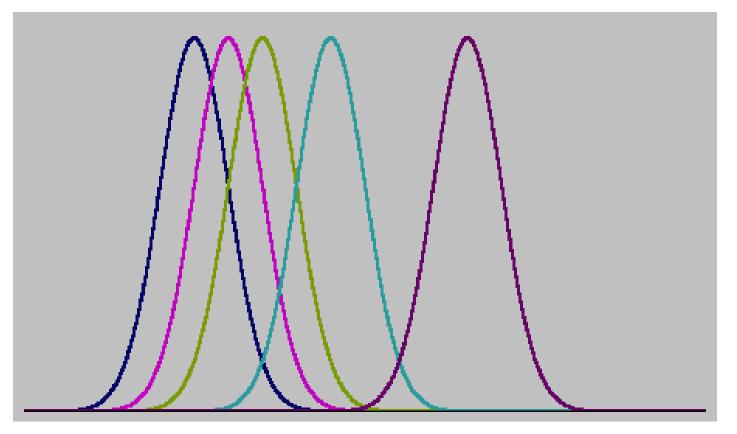
 Distribuciones normales con distintas desviación estándar e igual media.





#### LA MEDIA ( $\mu$ )

 Distribuciones normales con diferentes medias e igual desviación estándar.



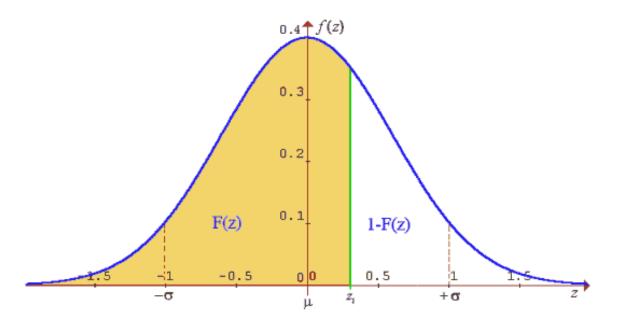


- Podemos concluir que en una distribución normal:
  - La desviación estándar ( $\sigma$ ) determina el grado de aplanamiento de la curva.
  - Cuanto mayor sea el valor de  $\sigma$ , más se dispersarán los datos en torno a la media y la curva será más plana.
  - La media indica la posición de la campana, de modo que para diferentes valores de  $\mu$  la gráfica es desplazada a lo largo del eje horizontal.
- La distribución normal que utilizaremos es la distribución normal estándar, que corresponde a una distribución de media 0 y varianza 1.

- Se le denomina como Z a la variable tipificada de X, y a la curva de su función de probabilidad se le conoce como la curva normal estándar.
- Es una distribución normal con promedio 0 y una desviación estándar de 1.
- Todas las variables normalmente distribuidas se pueden transformar a la distribución normal estándar utilizando la fórmula para calcular el valor Z correspondiente.
- Podemos decir que el valor de Z es la cantidad de desviaciones estándar a la que está distanciada la variable X del promedio.

# LA FUNCIÓN F(Z)

• En la siguiente gráfica vemos la representación gráfica de la función de *Z*.





# CARACTERÍSTICAS DE LA DISTRIBUCIÓN NORMAL ESTÁNDAR

- No depende de ningún parámetro.
- Su media es 0, su varianza es 1 y su desviación estándar es 1.
- La curva *f*(*x*) es simétrica respecto del eje de *Y*.
- Tiene un máximo en el eje de Y.
- Tiene dos puntos de inflexión en z=1 y z=-1.



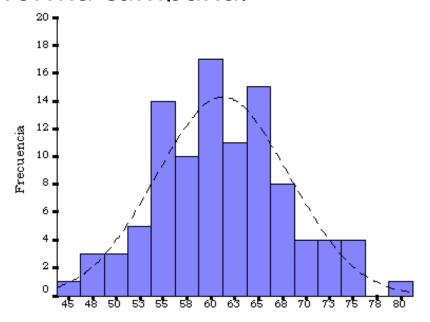
#### TEOREMA DEL LÍMITE CENTRAL

- Nos indica que, bajo condiciones muy generales, según aumenta la cantidad de datos, la distribución de la suma de variables aleatorias tenderá a seguir hacia una distribución normal.
- En otras palabras, el Teorema del Límite Central garantiza una distribución normal cuando el tamaño de la muestra es suficientemente grande.



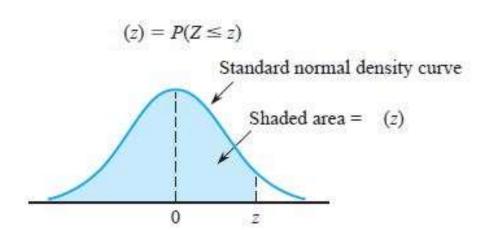
#### POR EJEMPLO:

- En el siguiente histograma podemos observar la distribución de frecuencias por peso de acuerdo a la edad.
- De acuerdo a este teorema según aumenten la cantidad de datos se podrá trazar una curva que tome cada vez más formación en forma campana.





- El área bajo la curva normal estándar es útil para asignar probabilidades de ocurrencia de la variable X.
- Debemos tomar en cuenta que el área total bajo la curva es igual a 1.
- Y que, por ser una gráfica simétrica, cada mitad tiene un área de 0,5.





2	.00	.01	.02	.03	.04	_05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	6064	.6103	.6141
0.3	6179	.6217	.6255	,6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	6985	.7019	.7054	.7088	7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	8078	.8106	.8133
0.9	.8159	.3186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	,8708	.8729	8749	.8770	8790	.8810	8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	,9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	9192	.9207	9222	.9236	.9251	.9265	.9278	.9292	.9306	_9319
1.5	.9332	.9345	.9357	.9370	.9382	9394	9406	.9418	9429	9441
1.6	.9452	.9463	9474	.9484	.9495	9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	9656	.9664	.9671	.9678	.9686	.9693	9699	.9706
1.9	.9713	.9719	9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	9783	.9788	9793	.9798	.9803	9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	9884	.9887	9890
2.3	.9893	9896	9898	.9901	.9904	.9906	9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	9948	.9949	.9951	.9952
2.6	.9953	.9955	9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	9974
2.8	.9974	.9975	9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	9990
3.1	.9990	.9991	9991	9991	9992	9992	9992	9992	.9993	9993
3.2	.9993	.9993	9994	9994	9994	9994	9994	9995	.9995	9995
3.3	.9995	.9995	9995	.9996	.9996	.9996	9995	9996	.9996	.9997
3.4	.9997	.9997	.9997	9997	.9997	.9997	.9997	9997	.9997	.9998



# PASOS PARA DETERMINAR EL ÁREA BAJO LA CURVA NORMAL ESTÁNDAR

- Paso 1 Interpretar gráficamente el área de interés.
- Paso 2 Determinar el valor Z.
- Paso 3 Buscar en la tabla de probabilidades.
- Paso 4 Hacer la suma o resta de áreas para encontrar la probabilidad deseada.



• La distancia entre un valor seleccionado, designado como X, y la población media  $\mu$ , dividida entre la desviación estándar de la población  $\sigma$ ,

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



- El ingreso mensual que una gran corporación ofrece a los graduados en MBA tiene una distribución normal con media de \$2000 y desviación estándar de \$200. ¿Cuál es el valor Z para un ingreso de \$2200? y ¿cuál para uno de \$1700?
  - Para X = \$2200, Z = (2200 2000) / 200 = 1.
  - Para X = \$1700, Z = (1700 2000) / 200 = -1,5.
  - Un valor Z igual a 1 indica que el valor de \$2200 es mayor que la desviación estándar de la media de \$2000.
  - Así como el valor Z igual a -1,5 indica que el valor de \$1700 es menor que la desviación estándar de la media de \$2000.



# ÁREAS BAJO LA CURVA NORMAL

PARAGUAYO ALEMANA

 Cerca de 68% del área bajo la curva normal está a más menos de una desviación estándar respecto a la media.

$$\mu + 1\sigma$$

 Alrededor de 95% está a más menos de dos desviaciones estándar de la media.

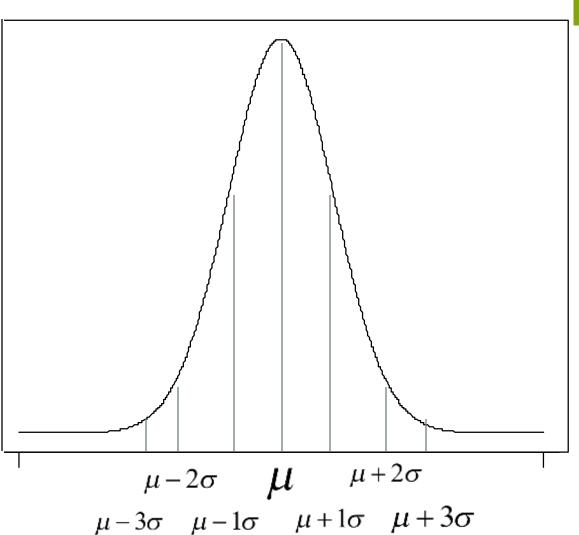
$$\mu + 2\sigma$$

 99,74% está a más menos de tres desviaciones estándar de la media.

$$\mu + 3\sigma$$



# ÁREAS BAJO LA CURVA NORMAL



PARAGUAYO ALEMANA

#### Entre:

1.68,26%

2.95,44%

3.99,74%



- El consumo de agua diario por persona en New Providence, Nueva Jersey tiene una distribución normal con media de 20 galones y desviación estándar de 5 galones.
  - Cerca de 68% del consumo de agua diario por persona en New Providence está entre cuáles dos valores.

$$\mu \pm 1\sigma = 20 \pm 1(5)$$
.

 Esto es, cerca de 68% del consumo diario de agua está entre 15 y 25 galones.



- ¿Cuál es la probabilidad de que una persona de New Providence seleccionada al azar use menos de 20 galones por día?
  - El valor Z asociado es:

$$Z = (20 - 20) / 5 = 0.$$
  
P(X<20) = P(Z<0) = 0,5

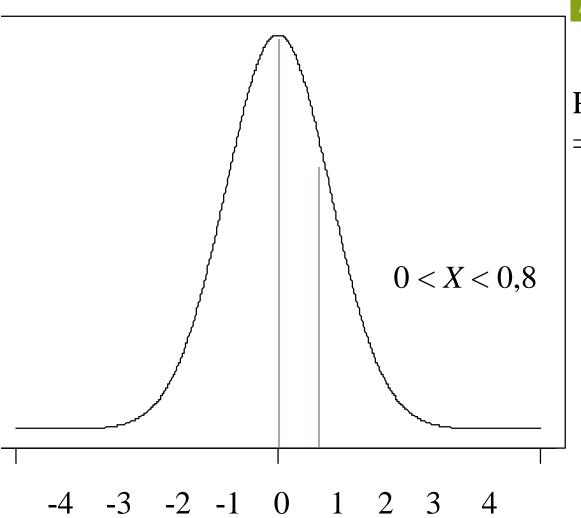
- ¿Qué porcentaje usan entre 20 y 24 galones?
- El valor z asociado con X = 20 es Z = 0 y con X = 24:

$$Z = (24 - 20) / 5 = 0.8$$
.  
P(X<24) = P(Z<0.8) = 0.7881

• Por lo tanto, P(20 < X < 24) = P(0 < Z < 0.8)0.7881 - 0.5 = 0.2881 = 28.81%



### **EJEMPLO 3**



$$P(0 < z < 0.8)$$
  
= 0.2881



#### **EJEMPLO 3**

- ¿Qué porcentaje de la población utiliza entre 18 y 26 galones?
  - El valor Z asociado con X = 18 es Z = (18 -20) /5 = -0.4, y para X = 26, Z = (26 20) /5 = 1,2.
  - Así, P(18 < X < 26) = P(-0.4 < Z < 1.2) = 0.8849 0.3458 = 0.5403

https://stattrek.com/online-calculator/normal.aspx



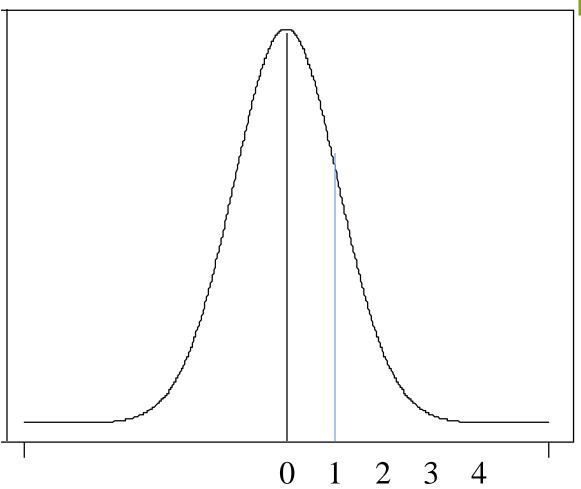
- El profesor Mann determinó que el promedio final en su curso de estadística tiene una distribución normal con media de 72 y desviación estándar de 5.
- Decidió asignar las calificaciones del curso de manera que 15% de los alumnos reciban una calificación de 5.
- ¿Cuál es el promedio más bajo que un alumno puede tener para obtener un 5?
  - Sea X el promedio más bajo. Encuentre X de manera que P(X > X) = 0.15 = 1-0.15 = 0.85
  - El valor *Z* correspondiente es 1,04.
  - Así se tiene (X 72) / 5 = 1,04, o X = 77,2



http://onlinestatbook.com/2/calculators/inverse\_normal\_dist.html

# **EJEMPLO 4**







- La cantidad de propina que un mesero recibe por turno en un restaurante exclusivo tiene una distribución normal con media de \$80 y desviación estándar de \$10.
- Shelli siente que ha dado un mal servicio si el total de sus propinas del turno es menor que \$65.
- ¿Cuál es la probabilidad de que ella haya dado un mal servicio?
  - Sea X la cantidad de propina.
  - El valor z asociado con X = 65 es Z = (65 80) / 10 = -1,5.
  - Así P(X<65) = P(Z<-1,5) = 1 (Z<1,5) = 1 0,9332 = 0,0668.



#### **EJEMPLO 6**

Supongamos ahora que  $X \sim N(100,16)$ .

a) ¿Cuál es la probabilidad de que la variable X tome un valor entre 100 y
 115? :

$$P(100 < X < 115) = P\left(\frac{100 - 100}{16} < \frac{X - \mu}{\sigma} < \frac{115 - 100}{16}\right) = P(0 < Z < 0.9375) \approx$$

$$\approx P(Z < 0.94) - P(Z < 0) = 0.8264 - 0.5000 = 0.3264$$

b) ¿Cúal es la probabilidad de que X tome un valor mayor de 90? :

$$P(X>90) = P\left(\frac{X-\mu}{\sigma} > \frac{90-100}{16}\right) = P(Z>-0.63) = 1 - P(Z<-0.63) = 1 - 0.2643 = 0.7357$$

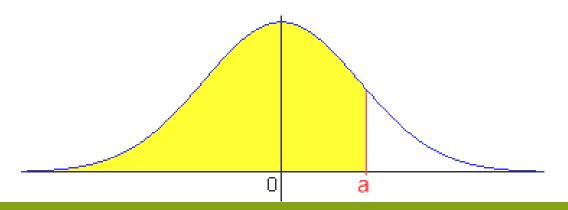


# Distribuciones de probabilidad de variables continuas

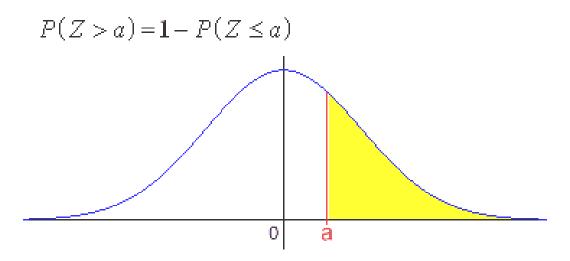
Ejercicios

- 1. Supongamos que sabemos que el peso de los/as estudiantes universitarios/as sigue una distribución aproximadamente normal, con una media de 140 libras y una desviación estándar de 20 libras.
- a) Determine la probabilidad de que una persona tenga un peso menor o igual a 150 libras

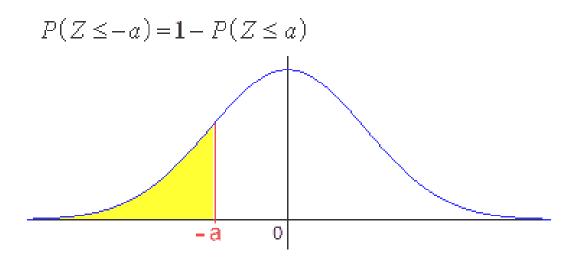
$$P(Z \le a) \to Tablas$$



2. Si deseamos la probabilidad de que una persona, elegida al azar, tenga un peso mayor o igual a 150 libras.

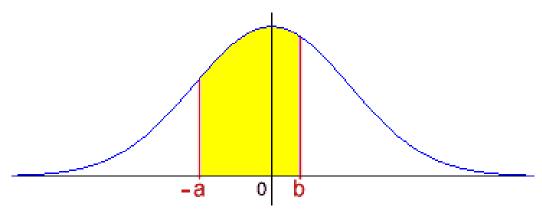


3. Determine la probabilidad de que una persona, elegida al azar, tenga un peso menor o igual a 115 libras.



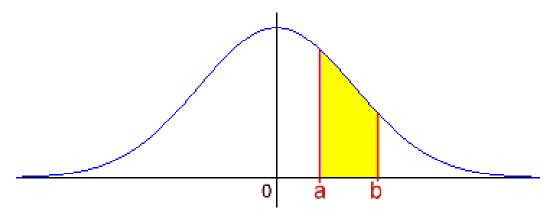
4. Si deseamos la probabilidad de que una persona, elegida al azar, tenga un peso entre 115 y 150 libras.

$$P(-a < Z \le b) = P(Z \le b) - [1 - P(Z \le a)]$$



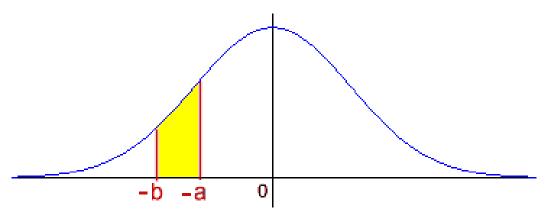
5. Si deseamos la probabilidad de que una persona, elegida al azar, tenga un peso entre 150 y 160 libras.

$$P(a < Z \le b) = P(Z \le b) - P(Z \le a)$$



6. Determine la probabilidad de elegir a una persona que pese entre 115 y 130 libras.

$$P(-b < Z \le -a) = P(a < Z \le b)$$



- 7. Sea Z una variable aleatoria normal estándar.
  - a) Hallar P(Z < 1,20)
  - b) Hallar P(Z > 1,33)
  - c) Hallar P(Z < -1,70)
  - d) Hallar P(Z > -1,00)
  - e) Hallar P (1.20 < Z< 1,33)
  - f) Hallar P (-1,70 < Z < 1,20)
  - g) Hallar P(-1,70 < Z < -1,00)

- 8. Una compañía produce un compuesto químico y está preocupada por su contenido de impurezas. Se estima que el peso de las impurezas por lote se distribuye según una normal con media 12,2 gramos y desviación típica 2,8 gramos. Se elige un lote al azar.
  - a) ¿Cuál es la probabilidad de que contenga menos de 10 gramos de impurezas?
  - b) ¿Cuál es la probabilidad de que contenga más de 15 gramos de impurezas?
  - c) ¿Cuál es la probabilidad de que contenga entre 12 y 15 gramos de impurezas?

PARAGUAYO ALEMANA

# ¡GRACIAS POR LA ATENCIÓN! eladio.martinez@upa.edu.py

