TECNICATURA SUPERIOR EN

**Desarrollo de Software**

## Probabilidad y Estadística Aplicada

# Unidad VI

# ÍNDICE

## Distribución de probabilidad

La distribución de probabilidad es una manera de relacionar situaciones que no pasaron, con medidas de tendencia central y variación. Para ello vamos a iniciar repasando algunas definiciones.

- Una variable aleatoria es aquella (casi siempre representada por x) que tiene un solo valor numérico determinado por el azar, para cada resultado de un procedimiento.

Una distribución de probabilidad es una distribución que indica la probabilidad de cada valor de la variable aleatoria. A menudo se expresa como gráfica, tabla o fórmula.

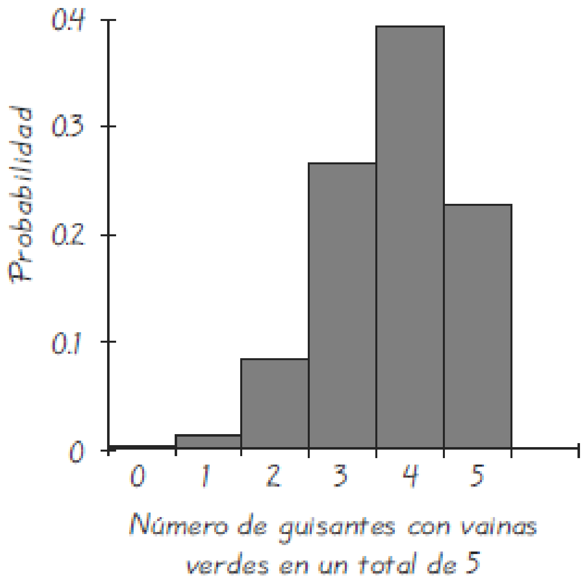
Ejemplo:

| $x$ (número de guisantes con vainas verdes) | $P(x)$ |
|---|---|
| 0 | 0.001 |
| 1 | 0.015 |
| 2 | 0.088 |
| 3 | 0.264 |
| 4 | 0.396 |
| 5 | 0.237 |

- Una **variable aleatoria discreta** tiene un número finito de valores o un número de valores contable, donde "contable" se refiere al hecho de que podría haber un número infinito de valores, pero que pueden asociarse con un proceso de conteo, de manera que el número de valores es 0 o 1 o 2 o 3, etcétera.
- Una **variable aleatoria continua** tiene un número infinito de valores, y esos valores pueden asociarse con mediciones en una escala continua, de manera que no existan huecos o interrupciones.

## Gráficas

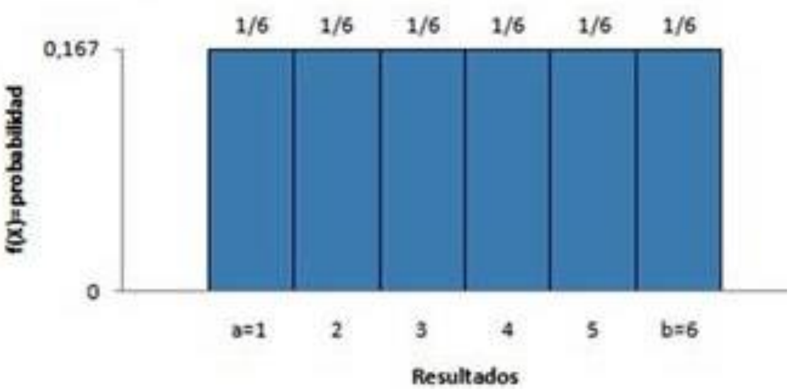
Existen varias formas para graficar una distribución de probabilidad, pero aquí consideraremos solamente el histograma de probabilidad, un histograma de probabilidad muy similar al histograma de frecuencias relativas pero la escala vertical indica probabilidades en vez de frecuencias relativas basadas en resultados muestrales reales.




Ejemplo: Lanzo un dado, Variable aleatoria discreta.

Espacio muestral: {1, 2, 3, 4, 5, 6}

Gráfica:



El histograma se realiza en base a una tabla de distribución de probabilidad

| x | P(x) |
|---|---|
| 1 | 1/6 |
| 2 | 1/6 |
| 3 | 1/6 |
| 4 | 1/6 |
| 5 | 1/6 |
| 6 | 1/6 |

## Requisitos de una distribución de probabilidad

1. $\sum P(x) = 1$ donde $x$ asume todos los valores posibles. (La suma de todas las probabilidades debe ser 1, pero valores como 0.999 o 1.001 son aceptables porque son el resultado de errores de redondeo).
2. $0 \leq P(x) \leq 1$ para cada valor individual de $x$. (Es decir, cada valor de probabilidad debe ubicarse entre 0 y 1, inclusive).

El primer requisito surge del simple hecho de que la variable aleatoria *x* representa todos los eventos posibles en el espacio muestral completo, de manera que tenemos la certeza (con probabilidad 1) de que uno de los eventos ocurrirá. En la tabla del ejemplo anterior, se observa que la suma de todas las probabilidades es 1, y que cada valor *P*(*x*) está entre 0 y 1. Puesto que la tabla satisface ambos requisitos, se confirma que se trata de una distribución de probabilidad. Una distribución de probabilidad puede describirse como una tabla una gráfica o una fórmula.

Ejemplo: Teléfonos celulares Según una encuesta realizada por Frank N. Magid Associates, la tabla:

| *x* | *P*(*x*) |
|---|---|
| 0 | 0.19 |
| 1 | 0.26 |
| 2 | 0.33 |
| 3 | 0.13 |

describe las probabilidades del número de teléfonos celulares en uso por familia. ¿esta tabla describe una distribución de probabilidad?

Para ser una distribución de probabilidad, $P(x)$ debe satisfacer los dos requisitos anteriores. Pero

$$\begin{aligned} \Sigma P(x) &= P(0) + P(1) + P(2) + P(3) \\ &= 0.19 + 0.26 + 0.33 + 0.13 \\ &= 0.91 \quad [\text{lo que demuestra que } \Sigma P(x) \neq 1] \end{aligned}$$

Como no se satisface el primer requisito, concluimos que la tabla *no* describe una distribución de probabilidad.

*Recordamos el cálculo de Media, varianza y desviación estándar:*

## Media, varianza y desviación estándar

| | |
|---|---|
| $\mu = \Sigma[x \cdot P(x)]$ | Media de una distribución de probabilidad |
| $\sigma^2 = \Sigma[(x - \mu)^2 \cdot P(x)]$ | Varianza de una distribución de probabilidad (más fácil de entender) |
| $\sigma^2 = \Sigma[x^2 \cdot P(x)] - \mu^2$ | Varianza de una distribución de probabilidad (cálculos más sencillos) |
| $\sigma = \sqrt{\Sigma[x^2 \cdot P(x)] - \mu^2}$ | Desviación estándar de una distribución de probabilidad |

Ejemplo:

**Cálculo de la media, la varianza y la desviación estándar**

En la tabla que se describe la distribución de probabilidad del número de guisantes con vainas verdes (Pág.3), en un total de 5 vástagos obtenidos de progenitores con pares de genes verde/amarillo.

Media: $\mu = \Sigma[x \cdot P(x)] = 3.752 = 3.8$ (se redondea)

Varianza: $\sigma^2 = \Sigma[(x - \mu)^2 \cdot P(x)] = 0.940574 = 0.9$ (se redondea)

La desviación estándar es la raíz cuadrada de la varianza; por lo tanto,

Desviación estándar: $\sigma = \sqrt{0.940574} = 0.969832 = 1.0$ (se redondea)

## Identificación de resultados Inusuales

Regla práctica de las desviaciones

$$\textbf{valor máximo común} = \mu + 2\sigma$$

$$\textbf{valor mínimo común} = \mu - 2\sigma$$

Si, con un supuesto dado, la probabilidad de un evento particular observado, es extremadamente pequeña, concluimos que el supuesto probablemente no es correcto.

Las probabilidades se pueden utilizar para aplicar la regla del evento inusual de la siguiente manera: Uso de las probabilidades para determinar resultados inusuales

• Número inusualmente alto de éxitos: x éxitos en n ensayos es un número inusualmente alto de éxitos si la probabilidad de x o **más** éxitos es poco probable, con una probabilidad de 0.05 o menor. Este criterio se puede expresar de la siguiente forma: P(x o más) ≤ 0.05.*

• Número inusualmente bajo de éxitos: x éxitos en n ensayos es un número inusualmente bajo de éxitos si la probabilidad de x o **menos** éxitos es poco probable, con una probabilidad de 0.05 o menor. Este criterio se puede expresar de la siguiente forma: P(x o menos) 0.05.*

(*El valor de 0.05 no es absolutamente rígido. Se podrían usar otros valores, como 0.01, para distinguir entre resultados que pueden ocurrir con facilidad por azar y eventos que tienen muy pocas probabilidades de ocurrir por azar.)

Ejemplo:

Identificación de resultados inusuales con la regla práctica de las desviaciones: En el ejemplo de los guisantes encontramos que, para grupos de 5 vástagos (generados a partir de progenitores con pares de genes de vainas verdes/amarillas), la media del número de guisantes con vainas verdes es 3.8 y la desviación estándar es 1.0.

Utilice esos resultados y la regla práctica de las desviaciones para calcular el valor máximo común y el valor mínimo común. Con base en los resultados, determine si es inusual que resulten 5 vástagos de guisantes y encontrar que solo uno de ellos tiene vainas verdes.

Al utilizar la regla práctica de las desviaciones, podemos calcular el valor máximo común y el valor mínimo común de la siguiente manera:

valor máximo común: $\mu + 2\sigma = 3.8 + 2(1.0) = 5.8$

valor mínimo común: $\mu - 2\sigma = 3.8 - 2(1.0) = 1.8$

Con base en estos resultados, concluimos que, para grupos de 5 vástagos de guisantes, el número de vástagos de guisantes con vainas verdes debe caer entre 1.8 y 5.8. Si resultan 5 vástagos de guisantes como los descritos, sería inusual obtener solo un vástago con vainas verdes (porque el valor de 1 está fuera de este rango de valores comunes: 1.8 a 5.8). (En este caso, el valor máximo común es en realidad 5, ya que es el número más grande posible de guisantes con vainas verdes).

## Valor esperado

La media de una variable aleatoria discreta es el resultado medio teórico de un número infinito de ensayos. Podemos considerar a esa media como el *valor esperado* en el sentido de que constituye el valor promedio que esperaríamos obtener si los ensayos pudieran continuar de manera indefinida. Los usos del valor esperado (también llamado *esperanza matemática* o simplemente *esperanza*) son extensos y variados, y desempeñan un papel muy importante en el área de aplicación denominada *teoría de la decisión*.

El **valor esperado** de una variable aleatoria discreta se denota con $E$ y representa el valor promedio de los resultados. Se obtiene calculando el valor de $\sum [x \, . \, P(x)]$

$$E = \sum [x \, . \, P(x)]$$

(No es necesario que un valor esperado sea un número entero, incluso si los diferentes valores posibles de $x$ son todos números enteros.)

Ejemplo:

Usted está considerando la posibilidad de apostar al número 7 de la ruleta o a la "línea de pase" en el juego de dados del casino Venetian de Las Vegas.

***a*)** Si usted apuesta $5 al número 7 de la ruleta, la probabilidad de perder $5 es 37/38, y la probabilidad de obtener una ganancia neta de $175 es 1/38. (El premio es de $180, incluyendo su apuesta de $5, por lo que la ganancia neta sería de $175). Calcule su valor esperado en caso de apostar $5 al número 7 de la ruleta.

***b*)** Si usted apuesta $5 a la línea de pase en los dados, la probabilidad de perder $5 es 251/495, y la probabilidad de obtener una ganancia neta de $5 es 244/495. (Si usted apuesta $5 a la línea de pase y gana, recibe $10 que incluyen su apuesta, de manera que la ganancia neta es de $5). Calcule su valor esperado si apuesta $5 a la línea de pase.

Qué es mejor: ¿una apuesta de $5 al número 7 en la ruleta, o una apuesta de $5 a la línea de pase en los dados? ¿Por qué?

Solución:

Ruleta: En la tabla se resumen las probabilidades y la ganancia al apostar $5 al número 7 en la ruleta. En esta tabla también se observa que el valor esperado es $\sum[x. \, P(x)]= -26$ centavos. Es decir, por cada apuesta de $5 al número 7, se espera perder un promedio de 26 centavos.

| Evento | $x$ | $P(x)$ | $x \cdot P(x)$ |
|---|---|---|---|
| Pérdida | −$5 | 37/38 | −$4.87 |
| Ganancia (neta) | $175 | 1/38 | $4.61 |
| Total | | | −$0.26 (o −26 centavos) |

Dados: En esta tabla se resumen las probabilidades y la ganancia al apostar $5 a la línea de pase en los dados. En esta tabla también se observa que el valor esperado es ∑[x. P(x)]=-8 centavos. Es decir, por cada apuesta de $5 a la línea de pase, se espera perder un promedio de 8 centavos.

| Evento | $x$ | $P(x)$ | $x \cdot P(x)$ |
|---|---|---|---|
| Pérdida | −$5 | 251/495 | −$2.54 |
| Ganancia (neta) | $5 | 244/495 | $2.46 |
| Total | | | −$0.08 (o −8 centavos) |

La apuesta de $5 en la ruleta produce un valor esperado de -26 centavos, y la apuesta de $5 en los dados produce un valor esperado de -8 centavos.
Es mejor apostar a los dados, ya que su valor esperado es mayor. Es decir, es mejor perder 8 centavos que 26. Aun cuando el juego de la ruleta ofrece la oportunidad de una mayor ganancia, a la larga, es mejor el juego de dados.

## Actividades:

1. Valor esperado Un investigador calcula el valor esperado del número de niñas en tres
nacimientos y obtiene un resultado de 1.5. Luego, redondea los resultados a 2, al afirmar que no es
posible que nazcan 1.5 niñas en cinco alumbramientos. ¿Es correcto este razonamiento? Explique.
2. Distribución de probabilidad Uno de los requisitos de una distribución de probabilidad
es que la suma de las probabilidades debe ser 1 (se permite una pequeña cantidad de variación por
errores de redondeo). ¿Cuál es la justificación de este requisito?
3. Distribución de probabilidad Un jugador profesional afirma que cargó un dado para que
los resultados de 1, 2, 3, 4, 5 y 6 tengan probabilidades correspondientes de 0.1, 0.2, 0.3, 0.4, 0.5
y 0.6. ¿Realmente será cierto lo que dice? ¿Una distribución de probabilidad se describe haciendo
una lista de los resultados junto con sus probabilidades correspondientes?
4) *identifique si la variable aleatoria que se describe es discreta o continua.*
5. *a*) El número de personas que ahora están conduciendo un automóvil en Estados Unidos

*b*) El peso del oro almacenado en Fort Knox

*c*) La altura del último avión que salió del aeropuerto JFK en la ciudad de Nueva York

*d*) El número de automóviles que chocaron el año pasado en San Francisco

*e*) El tiempo necesario para volar de Los Ángeles a Shangai

6. *a*) La cantidad total (en onzas) de bebidas gaseosas que usted consumió el año pasado

*b*) El número de latas de bebidas gaseosas que consumió el año pasado

*c*) El número de películas que actualmente se exhiben en los cines estadounidenses

*d*) La duración de una película elegida al azar

*e*) El costo de filmar una película elegida al azar

5) Identificación de distribuciones de probabilidad: *determine*

*si se trata o no de una distribución de probabilidad. Si se describe una distribución*

*de probabilidad, calcule su media y desviación estándar.*

a) Trastorno genético Tres hombres tienen un trastorno genético relacionado con el cromosoma X, y cada uno engendra un hijo. La variable aleatoria *x* es el número de hijos de los tres hombres que heredan el trastorno genético relacionado con el cromosoma X.

| $x$ | $P(x)$ |
|---|---|
| 0 | 0.125 |
| 1 | 0.375 |
| 2 | 0.375 |
| 3 | 0.125 |

b) Vuelos sobrevendidos Air America tiene la política de sobrevender sus vuelos de forma habitual. La variable aleatoria *x* representa el número de pasajeros que no pueden abordar debido a que hay más pasajeros que asientos (según datos de un documento de investigación de IBM, escrito por Lawrence, Hong y Cherrier).

| $x$ | $P(x)$ |
|---|---|
| 0 | 0.051 |
| 1 | 0.141 |
| 2 | 0.274 |
| 3 | 0.331 |
| 4 | 0.187 |

## Distribución de probabilidad Binomial

La distribución de probabilidad binomial nos permite enfrentar circunstancias donde los resultados pertenecen a dos categorías relevantes, tales como Aceptable/defectuoso o sobrevivió/murió.

Una distribución de probabilidad binomial resulta de un procedimiento que cumple con todos los siguientes requisitos:

1. El procedimiento tiene un *número fijo de ensayos*.
2. Los ensayos deben ser *independientes*. (El resultado de cualquier ensayo individual no afecta las probabilidades de los demás ensayos).
3. Todos los resultados de cada ensayo deben clasificarse en *dos categorías* (generalmente llamadas *éxito* y *fracaso*).
4. La probabilidad de un éxito permanece igual en todos los ensayos.

Requisito de independencia Al seleccionar una muestra (por ejemplo, sujetos de encuestas) para un análisis estadístico, generalmente se hace el muestreo sin reemplazo.

Recuerde que el muestreo sin reemplazo implica eventos dependientes, lo cual viola el segundo requisito de la definición anterior. Sin embargo, a menudo es posible suponer independencia al aplicar el siguiente lineamiento del 5%.

Tratar eventos dependientes como independientes: El lineamiento del 5% para cálculos engorrosos.

Si los cálculos son engorrosos y el tamaño de muestra no es mayor que el 5% del tamaño de la población, trate las selecciones como independientes (incluso si las selecciones se efectúan sin reemplazo, de modo que sean técnicamente dependientes).

Si un procedimiento satisface los cuatro requisitos mencionados, la distribución de la variable aleatoria $x$ (número de éxitos) se denomina *distribución de probabilidad binomial* (o *distribución binomial*), en la que suele usarse la siguiente notación.

*Notación:*

| **Notación para distribuciones de probabilidad binomial** | |
|---|---|
| E y F (éxito y fracaso) denotan las dos categorías posibles de todos los resultados. | |
| $P(E) = p$ | ($p$ = probabilidad de un éxito) |
| $P(F) = 1 - p = q$ | ($q$ = probabilidad de un fracaso) |
| $n$ | denota el número fijo de ensayos. |
| $x$ | denota un número específico de éxitos en $n$ ensayos, de manera que $x$ puede ser cualquier número entero entre 0 y $n$, inclusive. |
| $p$ | denota la probabilidad de *éxito* en *uno* de $n$ ensayos. |
| $q$ | denota la probabilidad de *fracaso* en *uno* de $n$ ensayos. |
| $P(x)$ | denota la probabilidad de lograr exactamente $x$ éxitos en los $n$ ensayos. |

*Ejemplo:*

Considere un experimento en el que se generan 5 vástagos de guisantes de plantas progenitoras que tienen la combinación de genes verde/amarillo para el color de las vainas. En el problema del capítulo vimos que la probabilidad de que un vástago de guisantes tenga vainas verdes es 3/4 o 0.75. Es decir, *P* (vaina verde) = 0.75. Suponga que queremos calcular la probabilidad de que exactamente 3 de los 5 vástagos tengan vainas verdes.

*a*) ¿Este procedimiento da como resultado una distribución binomial?

*b*) Si este procedimiento da como resultado una distribución binomial, identifique los valores de $n$, $x$, $p$ y $q$.

Este procedimiento satisface los requisitos de una distribución binomial, como se muestra a continuación.

1. El número de ensayos (5) es fijo.
2. Los 5 ensayos son independientes, ya que la probabilidad de que cualquier vástago tenga vainas verdes no se ve afectada por el resultado de cualquier otro vástago.
3. Cada uno de los 5 ensayos tiene dos categorías de resultados: el guisante tiene vainas verdes o no las tiene.
4. Para cada vástago de guisante, la probabilidad de que tenga vainas verdes es ¾ o 0.75, y la probabilidad es la misma para los 5 guisantes.

*b*) Luego de haber concluido que el procedimiento da como resultado una distribución binomial, ahora procederemos a identificar los valores de $n$, $x$, $p$ y $q$.

1. Con 5 vástagos de guisantes, tenemos $n = 5$.

2. Queremos conocer la probabilidad de exactamente 3 guisantes con vainas verdes,

de manera que $x = 3$.

3. La probabilidad de éxito (obtener un guisante con vainas verdes) en una selección es 0.75, de modo que $p = 0.75$.

4. La probabilidad de fracaso (no tener un guisante con vainas verdes) es 0.25, de modo que $q = 0.25$.

Nuevamente, es muy importante asegurarse de que $x$ y $p$ se refieran al mismo concepto de "éxito". En este ejemplo, se utiliza $x$ para contar el número de guisantes con vainas verdes, de modo que $p$ debe ser la probabilidad de que un guisante tenga vainas verdes.

Por consiguiente, $x$ y $p$ utilizan el mismo concepto de éxito aquí (vainas verdes).

**Uso de la fórmula de probabilidad binomial**

En una distribución de probabilidad binomial, las probabilidades pueden calcularse mediante la fórmula de probabilidad binomial:

$$P(x) = \frac{n!}{(n-x)!x!} \cdot p^x \cdot q^{n-x} \quad \text{para } x = 0, 1, 2, \ldots, n$$

donde
$n$ = número de ensayos
$x$ = número de éxitos en $n$ ensayos
$p$ = probabilidad de éxito en cualquier ensayo
$q$ = probabilidad de fracaso en cualquier ensayo ($q = 1 - p$)

*Ejemplo:*

Suponiendo que la probabilidad de que un guisante tenga vainas verdes es 0.75 (como en el problema del capítulo y en el ejemplo 1), utilice la fórmula de probabilidad binomial para calcular la probabilidad de obtener exactamente

3 guisantes con vainas verdes cuando se generan 5 vástagos de guisantes. Es decir, calcule $P(3)$ dado que $n = 5$, $x = 3$, $p = 0.75$ y $q = 0.25$.

Al emplear los valores dados de $n$, $x$, $p$ y $q$ en la fórmula de probabilidad binomial (fórmula), obtenemos

$$P(3) = \frac{5!}{(5-3)!\,3!} \cdot 0.75^3 \cdot 0.25^{5-3}$$
$$= \frac{5!}{2!3!} \cdot 0.421875 \cdot 0.0625$$
$$= (10)(0.421875)(0.0625) = 0.263671875$$

La probabilidad de obtener exactamente 3 guisantes con vainas verdes de un total de 5 vástagos es 0.264 (redondeado a tres dígitos significativos).

*Otra forma de calcular es utilizando herramientas tecnológicas: STATDISK, Minitab, Excel,SPSS, SAS y la calculadora TI-83/84 Plus pueden usarse para calcular probabilidades binomiales.*

## Distribución de probabilidad de Poisson

La **distribución de Poisson** es una distribución de probabilidad discreta que se aplica a las ocurrencias de algún evento *durante un intervalo específico.* La variable aleatoria *x* es el número de veces que ocurre un evento en un intervalo. El intervalo puede ser tiempo, distancia, área, volumen o alguna unidad similar. La probabilidad de que el evento ocurra *x* veces durante un intervalo está dada por la fórmula:

$$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!} \qquad \text{donde } e \approx 2.71828$$

**Requisitos de la distribución de Poisson**

**1.** La variable aleatoria *x* es el número de veces que ocurre un evento *durante un intervalo*.

**2.** Las ocurrencias deben ser *aleatorias*.

**3.** Las ocurrencias deben ser *independientes* entre sí.

**4.** Las ocurrencias deben estar *uniformemente distribuidas* dentro del intervalo considerado.

**Parámetros de la distribución de Poisson**

- La media es $\mu$.
- La desviación estándar es $\sigma = \sqrt{\mu}$.

La distribución de Poisson difiere de una distribución binomial en estas formas fundamentales:

**1.** La distribución binomial se ve afectada por el tamaño de la muestra *n* y la probabilidad

*p*, mientras que la distribución de Poisson solo se ve afectada por la media m.

**2.** En una distribución binomial, los valores posibles de la variable aleatoria *x* son 0, 1,…, *n*, pero los valores posibles *x* de una distribución de Poisson son 0, 1, 2,…, sin límite superior.

Ejemplo:

En un periodo reciente de 100 años, hubo 93 grandes terremotos (con una magnitud de al menos 6.0 en la escala de Richter) en el mundo (según datos de World Almanac and Book of Facts). Suponga que la distribución de Poisson es un modelo adecuado.

a) Calcule la media del número de grandes terremotos que ocurren cada año.

b) Si P(x) es la probabilidad de x terremotos en un año elegido al azar, calcule P(0), P(1), P(2), P(3), P(4), P(5), P(6) y P(7).

c) Los resultados reales son los siguientes: 47 años (0 grandes terremotos); 31 años (un gran terremoto); 13 años (2 grandes terremotos); 5 años (3 grandes terremotos); 2 años (4 grandes terremotos); 0 años (5 grandes terremotos); un año (6 grandes terremotos); 1 año (7 grandes terremotos). ¿Qué diferencia hay entre los resultados reales y las probabilidades obtenidas en el inciso *b*)? Al parecer, ¿la distribución de Poisson es un buen modelo en este caso?

Solución:

***a*)** Aplicamos la distribución de Poisson, ya que estamos tratando con las ocurrencias de

un evento (terremotos) dentro de un intervalo (un año). El número medio de terremotos

por año es

$$\mu = \frac{\text{número de terremotos}}{\text{número de años}} = \frac{93}{100} = 0.93$$

## Distribución Normal estándar

La *distribución normal estándar*, la cual tiene las siguientes tres propiedades:

**1.** Su gráfica tiene forma de campana.

**2.** Posee una media igual a 0.

**3.** Tiene una desviación estándar igual a 1.

**Distribuciones uniformes**

La distribución uniforme nos facilita visualizar estas dos propiedades muy importantes:

**1.** El área bajo la curva de una distribución de probabilidad es igual a 1.

**2.** Existe una correspondencia entre el área y la probabilidad (o frecuencia relativa), de manera que algunas probabilidades se pueden calcular al identificar las áreas correspondientes.

Una variable aleatoria continua tiene una **distribución uniforme** si sus valores se dispersan *uniformemente* a través del rango de posibilidades. La gráfica de una distribución
uniforme tiene forma rectangular.


La gráfica de una distribución de probabilidad continua, se llama **función de densidad,** y debe satisfacer los siguientes dos requisitos.

Requisitos de una función de densidad

**1.** El área total bajo la curva debe ser igual a 1.

**2.** Cada punto de la curva debe tener una altura vertical igual a o mayor que 0. (Es decir, la curva no puede estar por debajo del eje *x*).

**Distribución normal estándar**

La función de densidad de una distribución uniforme es una línea horizontal, de manera que es fácil calcular el área de cualquier región rectangular aplicando la siguiente fórmula:

Área = anchura * altura. Debido a que la función de densidad de una distribución normal

tiene una forma de campana más complicada, es más difícil calcular áreas, pero el principio básico es el mismo: *existe una correspondencia entre área y probabilidad*.

La **distribución normal estándar** es una distribución normal de probabilidad con media= 0 y desviación estándar =1, y el área total debajo de su función de densidad es igual a 1.

Como no es fácil calcular áreas, los matemáticos han calculado muchas áreas diferentes bajo la curva, las cuales se incluyen en la tabla.

Si utiliza la tabla , es esencial que comprenda lo siguiente:

**1.** La tabla está diseñada únicamente para la distribución normal *estándar*, que tiene una media de 0 y una desviación estándar de 1.

**2.** La tabla abarca dos páginas, una para las puntuaciones *z negativas* y la otra para las puntuaciones *z positivas*.

Cada valor en la tabla es un *área acumulada desde la izquierda* hasta un límite vertical por arriba de una puntuación *z* específica.

**4.** Cuando construya una gráfica, evite la confusión entre las puntuaciones *z* y las áreas.

**Puntuación *z*: *Distancia* a lo largo de la escala horizontal de la distribución normal estándar; remítase a la columna de la extrema izquierda y al renglón superior de la tabla**

**Área: *Región* bajo la curva; remítase a los valores de la tabla .**

**5.** La parte de la puntuación *z* que denota centésimas se encuentra en el renglón superior de la tabla

Ejemplo:

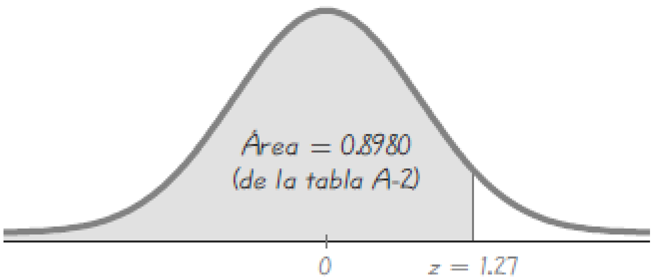
**Termómetros científicos** La Precision Scientific Instrument Company fabrica termómetros que, se supone, deben dar lecturas de 0°C en el punto de congelación del agua. Las pruebas de una muestra grande de estos instrumentos revelaron que, en el punto de congelación del agua, algunos termómetros daban lecturas por debajo de 0° (denotadas con números negativos), y otros daban lecturas por encima de 0° (denotadas con números positivos). Suponga que la lectura media es 0°C y que la desviación estándar de las lecturas es 1.00°C. También suponga que las lecturas se distribuyen de manera normal. Si se elige al azar un termómetro, calcule la probabilidad de que, en el punto de congelación del agua, la lectura sea menor que 1.27°.

Parte de la tabla:

| z | .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 |
| 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 | .9279 | .9292 | .9306 | .9319 |

Solución:

La distribución de probabilidad de las lecturas es una distribución normal estándar, ya que las lecturas se distribuyen de forma normal, con media= 0 y desviación estándar= 1. Necesitamos encontrar el área que está debajo de $z$ = 1.27. El *área* por debajo de $z$ =1.27 es igual a la *probabilidad* de seleccionar al azar un termómetro con una lectura menor que 1.27°. En la tabla encontramos que esta área es 0.8980.




La *probabilidad* de seleccionar al azar un termómetro con una lectura menor que 1.27° (en el punto de congelación del agua) es igual al área de 0.8980, que aparece como la región sombreada.

## Actividades:

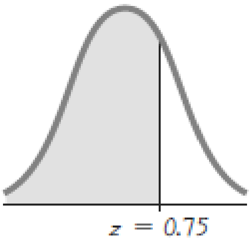
**1. Distribución normal** Cuando nos referimos a una distribución “normal”, ¿el término “normal” tiene el mismo significado que en el lenguaje cotidiano, o tiene un significado especial en estadística? ¿Qué es exactamente una distribución normal?

**2. Distribución normal** Una distribución normal se describe de manera informal como una distribución de probabilidad con “forma de campana” cuando se grafica. Describa la “forma de campana”.

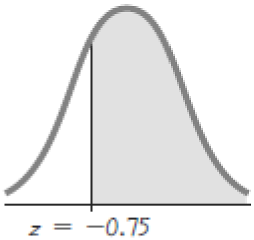
**3. Distribución normal estándar** ¿Qué requisitos son necesarios para que una distribución de probabilidad normal sea una distribución de probabilidad normal *estándar*?

4. ***La gráfica describe la distribución normal estándar con media igual a 0 y desviación estándar igual a 1***

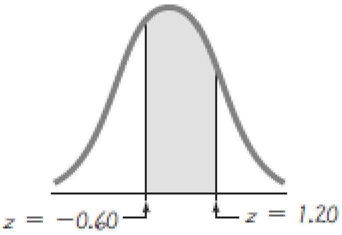
**9.**



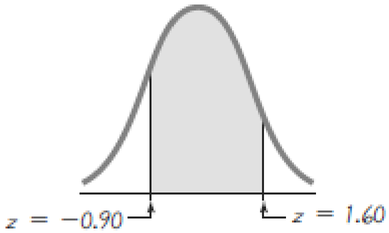

**10.**




**11.**




**12.**




## Bibliografía:


- **Estadística, Mario F. Triola, Decimoprimera edición**


## Anexo:

# Puntuaciones *z* NEGATIVAS

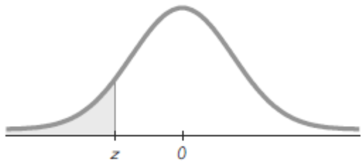


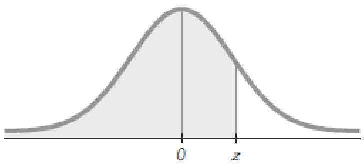

**TABLA A-2** Distribución normal estándar (*z*): Área acumulativa desde la IZQUIERDA

| *z* | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| −3.50 y menores | .0001 | | | | | | | | | |
| −3.4 | .0003 | .0003 | .0003 | .0003 | .0003 | .0003 | .0003 | .0003 | .0003 | .0002 |
| −3.3 | .0005 | .0005 | .0005 | .0004 | .0004 | .0004 | .0004 | .0004 | .0004 | .0003 |
| −3.2 | .0007 | .0007 | .0006 | .0006 | .0006 | .0006 | .0006 | .0005 | .0005 | .0005 |
| −3.1 | .0010 | .0009 | .0009 | .0009 | .0008 | .0008 | .0008 | .0008 | .0007 | .0007 |
| −3.0 | .0013 | .0013 | .0013 | .0012 | .0012 | .0011 | .0011 | .0011 | .0010 | .0010 |
| −2.9 | .0019 | .0018 | .0018 | .0017 | .0016 | .0016 | .0015 | .0015 | .0014 | .0014 |
| −2.8 | .0026 | .0025 | .0024 | .0023 | .0023 | .0022 | .0021 | .0021 | .0020 | .0019 |
| −2.7 | .0035 | .0034 | .0033 | .0032 | .0031 | .0030 | .0029 | .0028 | .0027 | .0026 |
| −2.6 | .0047 | .0045 | .0044 | .0043 | .0041 | .0040 | .0039 | .0038 | .0037 | .0036 |
| −2.5 | .0062 | .0060 | .0059 | .0057 | .0055 | .0054 | .0052 | .0051 * | .0049 | .0048 |
| −2.4 | .0082 | .0080 | .0078 | .0075 | .0073 | .0071 | .0069 | .0068 | .0066 | .0064 |
| −2.3 | .0107 | .0104 | .0102 | .0099 | .0096 | .0094 | .0091 | .0089 | .0087 | .0084 |
| −2.2 | .0139 | .0136 | .0132 | .0129 | .0125 | .0122 | .0119 | .0116 | .0113 | .0110 |
| −2.1 | .0179 | .0174 | .0170 | .0166 | .0162 | .0158 | .0154 | .0150 | .0146 | .0143 |
| −2.0 | .0228 | .0222 | .0217 | .0212 | .0207 | .0202 | .0197 | .0192 | .0188 | .0183 |
| −1.9 | .0287 | .0281 | .0274 | .0268 | .0262 | .0256 | .0250 | .0244 | .0239 | .0233 |
| −1.8 | .0359 | .0351 | .0344 | .0336 | .0329 | .0322 | .0314 | .0307 | .0301 | .0294 |
| −1.7 | .0446 | .0436 | .0427 | .0418 | .0409 | .0401 | .0392 | .0384 | .0375 | .0367 |
| −1.6 | .0548 | .0537 | .0526 | .0516 | .0505 * | .0495 | .0485 | .0475 | .0465 | .0455 |
| −1.5 | .0668 | .0655 | .0643 | .0630 | .0618 | .0606 | .0594 | .0582 | .0571 | .0559 |
| −1.4 | .0808 | .0793 | .0778 | .0764 | .0749 | .0735 | .0721 | .0708 | .0694 | .0681 |
| −1.3 | .0968 | .0951 | .0934 | .0918 | .0901 | .0885 | .0869 | .0853 | .0838 | .0823 |
| −1.2 | .1151 | .1131 | .1112 | .1093 | .1075 | .1056 | .1038 | .1020 | .1003 | .0985 |
| −1.1 | .1357 | .1335 | .1314 | .1292 | .1271 | .1251 | .1230 | .1210 | .1190 | .1170 |
| −1.0 | .1587 | .1562 | .1539 | .1515 | .1492 | .1469 | .1446 | .1423 | .1401 | .1379 |
| −0.9 | .1841 | .1814 | .1788 | .1762 | .1736 | .1711 | .1685 | .1660 | .1635 | .1611 |
| −0.8 | .2119 | .2090 | .2061 | .2033 | .2005 | .1977 | .1949 | .1922 | .1894 | .1867 |
| −0.7 | .2420 | .2389 | .2358 | .2327 | .2296 | .2266 | .2236 | .2206 | .2177 | .2148 |
| −0.6 | .2743 | .2709 | .2676 | .2643 | .2611 | .2578 | .2546 | .2514 | .2483 | .2451 |
| −0.5 | .3085 | .3050 | .3015 | .2981 | .2946 | .2912 | .2877 | .2843 | .2810 | .2776 |
| −0.4 | .3446 | .3409 | .3372 | .3336 | .3300 | .3264 | .3228 | .3192 | .3156 | .3121 |
| −0.3 | .3821 | .3783 | .3745 | .3707 | .3669 | .3632 | .3594 | .3557 | .3520 | .3483 |
| −0.2 | .4207 | .4168 | .4129 | .4090 | .4052 | .4013 | .3974 | .3936 | .3897 | .3859 |
| −0.1 | .4602 | .4562 | .4522 | .4483 | .4443 | .4404 | .4364 | .4325 | .4286 | .4247 |
| −0.0 | .5000 | .4960 | .4920 | .4880 | .4840 | .4801 | .4761 | .4721 | .4681 | .4641 |

*NOTA:* En el caso de valores de *z* por debajo de −3.49, utilice 0.0001 para el área.

*Utilice estos valores comunes que resultan de la interpolación:

| Puntuación *z* | Área |
|---|---|
| −1.645 | 0.0500 |
| −2.575 | 0.0050 |




# Puntuaciones *z* POSITIVAS

**TABLA A-2** (*continuación*) Área acumulativa desde la IZQUIERDA

| *z* | .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 | .8186 | .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 | .9279 | .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 | .9996 | .9996 | .9996 | .9997 |
| 3.4 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9998 |
| 3.50 y mayores | .9999 | | | | | | | | | |

*NOTA:* En el caso de valores de *z* por encima de 3.49, utilice 0.9999 para el área.

*Utilice estos valores comunes que resultan de la interpolación:

| Puntuación *z* | Área |
|---|---|
| 1.645 | 0.9500 |
| 2.575 | 0.9950 |

**Valores críticos comunes**

| Niveles de confianza | Valor crítico |
|---|---|
| 0.90 | 1.645 |
| 0.95 | 1.96 |
| 0.99 | 2.575 |