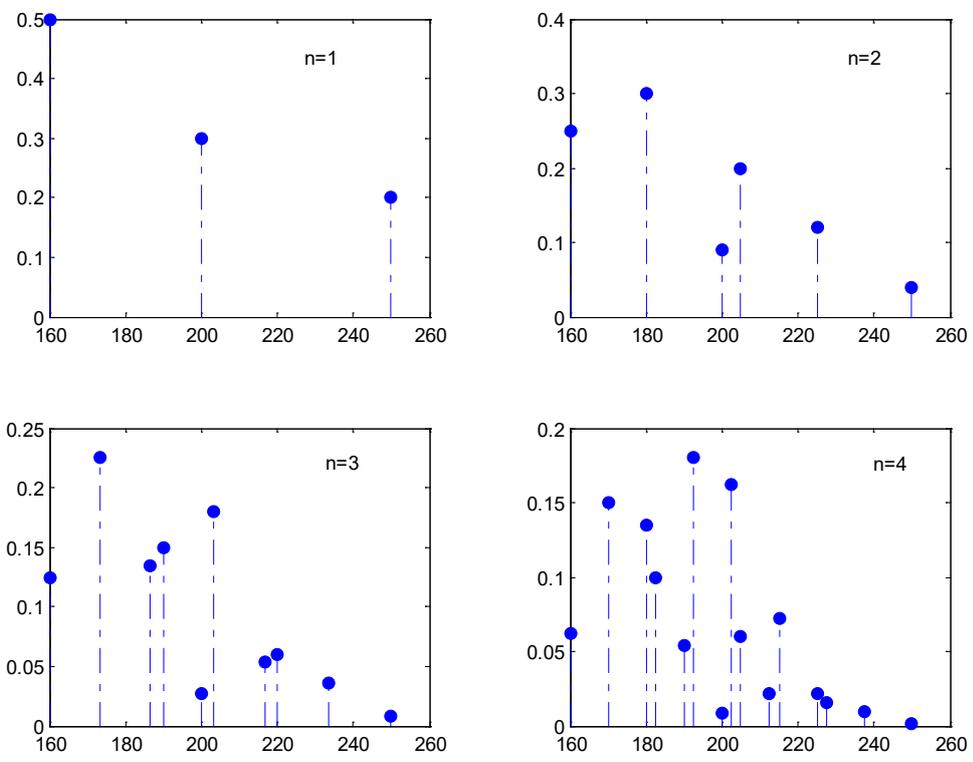


Trabajo Práctico 6. Distribución del muestreo y Teorema Central del Límite

Ejercicio 1. Un vendedor ofrece de un mismo producto, tres calidades diferentes. De todos los clientes que compran un sólo artículo, el 20% compra el de más alta calidad cuyo costo es de \$250, el 30% compra de calidad media cuyo costo es de \$200, y el 50% restante compra el de más baja calidad cuyo costo es de \$160. Sea X el pago efectuado por un cliente seleccionado al azar que se llevan un sólo artículo.

- a) Obtener la distribución de probabilidad puntual de X y hallar $E[X]=\mu$ y $V(X)=\sigma^2$.
- b) Se considera ahora una muestra aleatoria de tamaño 2, esto es el conjunto de dos variables aleatorias independientes e idénticamente distribuidas, X_1 y X_2 , correspondiente a los pagos efectuados por dos cliente. Determinar la distribución de probabilidad puntual del promedio muestral $\bar{X} = (X_1 + X_2)/2$. Calcular $E[\bar{X}]$ y comparar con μ .
- c) Calcular la probabilidad $P(\bar{X} > 210)$.
- d) Para la misma muestra aleatoria de tamaño 2, demostrar que la varianza muestral $S^2 = \frac{1}{2-1} \sum_{i=1}^2 (X_i - \bar{X})^2$ es igual a $S^2 = \frac{1}{2} (X_1 - X_2)^2$.
- e) Hallar la distribución de probabilidad puntual de la varianza muestral S^2 , calcular $E[S^2]$ y comparar con $V(X)=\sigma^2$.
- f) Constatar que los gráficos correspondientes a la distribución de probabilidad del promedio muestral $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, esto es el promedio de pagos efectuados por n clientes, para $n=1, 2, 3$ y 4 son los dados a continuación.

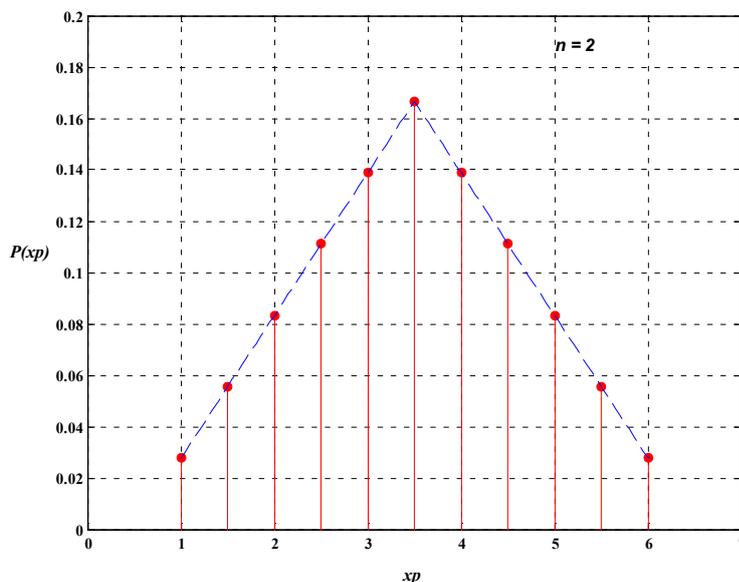


Ejercicio 2. Sea X la variable aleatoria discreta correspondiente al número que sale en la cara superior al arrojar un dado equilibrado.

- Obtener la distribución de probabilidad puntual de X y hallar $E[X]=\mu$ y $V(X)=\sigma^2$. Calcular $P(3 \leq X \leq 4)$.
- Considerar la enumeración de todas las muestras posibles de tamaño 2 correspondientes a la variable aleatoria X , esto es el conjunto de las dos variables aleatorias independientes e idénticamente distribuidas, X_1 y X_2 , correspondiente a los resultados que se obtienen al tirar dos veces el dado. Determinar la distribución de probabilidad puntual del promedio muestral $\bar{X} = (X_1 + X_2)/2$ y corroborar que la misma está definida por

$$P(\bar{X} = \bar{x}) = \frac{6 - \left| \bar{x} - \frac{7}{2} \right|}{36} \quad \bar{x} = \frac{k}{2} \wedge k = 2, 3, 4, \dots, 12,$$

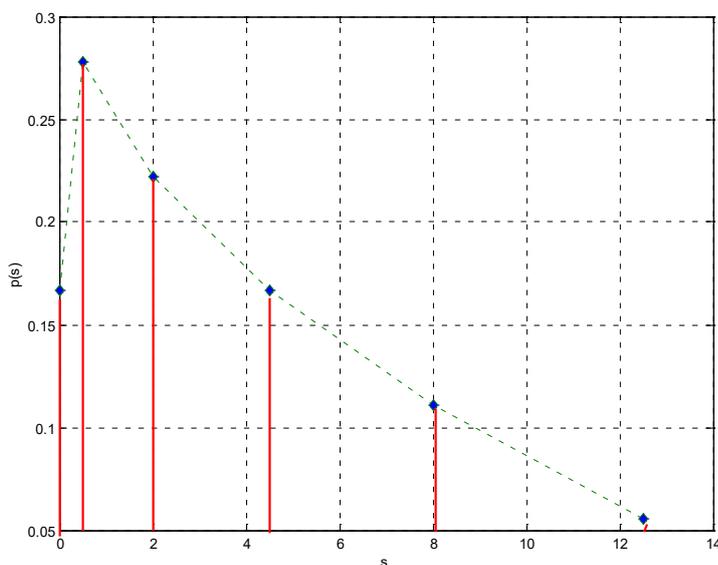
cuya representación gráfica es la dada en la figura *a*.



Ej.2. Fig a

Hallar $E[\bar{X}]$ y comparar con μ . Calcular $P(3 \leq \bar{X} \leq 4)$.

- Hallar la distribución de probabilidad puntual de la varianza muestral S^2 , calcular $E[S^2]$ y comparar con $V(X)=\sigma^2$. Constatar que la representación gráfica es la dada en la figura *b*.

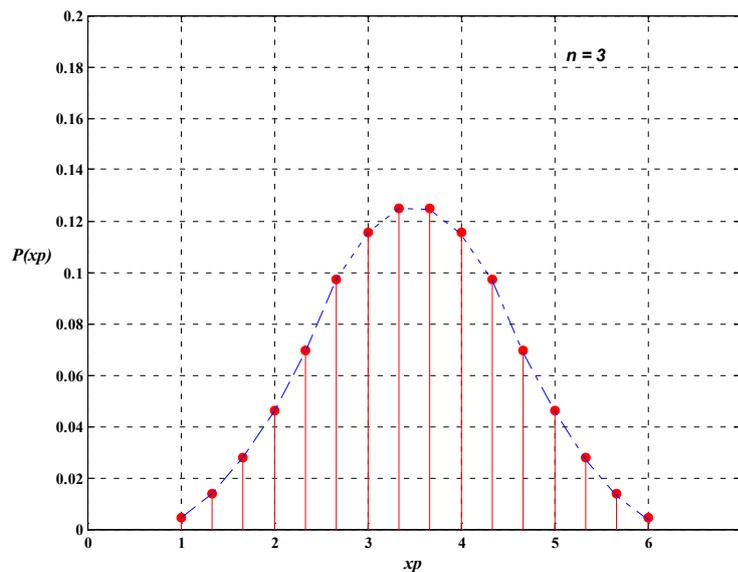


Ej.2. Fig b

- d) Considerar la enumeración de todas las muestras posibles de tamaño 3 correspondientes a la variable aleatoria X , esto es el conjunto de las dos variables aleatorias independientes e idénticamente distribuidas, X_1, X_2 y X_3 correspondiente a los resultados que se obtienen al tirar tres veces el dado. Determinar la distribución de probabilidad puntual del promedio muestral $\bar{X} = (X_1 + X_2 + X_3)/3$ y corroborar que la misma está definida por

$$\begin{cases} \frac{9}{432}\bar{x}^2 - \frac{9}{432}\bar{x} + \frac{1}{216} & 1 \leq \bar{x} < 3 \\ -\frac{9}{216}\bar{x}^2 + \frac{63}{216}\bar{x} - \frac{83}{216} & 3 \leq \bar{x} < 4 \text{ para } \bar{x} \in \mathbb{R}_{\bar{x}}, \\ \frac{9}{432}\bar{x}^2 + \frac{111}{432}\bar{x} - \frac{171}{216} & 4 \leq \bar{x} < 6 \end{cases}$$

cuya representación gráfica es la dada en la figura c.



Ej.2. Fig c

Hallar $E[\bar{X}]$ y comparar con μ . Calcular $P(3 \leq \bar{X} \leq 4)$.

Ejercicio 3. Sea X_1, X_2, \dots, X_n una muestra aleatoria (**conjunto de variables aleatorias idénticamente distribuidas e independientes**) de una variable aleatoria X con valor esperado μ y varianza σ^2 . Demostrar que la variable aleatoria varianza muestral definida por

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2,$$

donde \bar{X} es el promedio muestral satisface que $E[S^2] = \sigma^2$.

Ejercicio 4. Transporte en camiones. El peso X , en kg, de las cajas transportadas por la Camionera Fierrorrap se distribuye de manera **aproximadamente normal** con $E[X] = \mu = 20$ kg y $V(X) = \sigma^2 = 9$ kg².

Calcular la probabilidad de que:

- el peso de una caja tomada al azar esté comprendido entre 19,7 kg y 20,6 kg;
- el peso medio de una muestra de 100 cajas, $\bar{X} = \frac{X_1 + X_2 + \dots + X_{100}}{100}$, esté comprendido entre 19,7 kg y 20,6 kg;
- un camión que transporta 100 cajas sea multado si se conoce que dicha sanción se aplica si el peso neto de su carga, $T = X_1 + X_2 + \dots + X_{100}$, excede los 2060 kg.

- d) Determinar un intervalo simétrico con respecto a μ que abarque el peso medio de 100 de estas cajas con una probabilidad de 0,95. Ayuda: determinar el número real positivo a de forma que $P(\mu - a < \bar{X} < \mu + a) = 0,95$.

Ejercicio 5. Suma y Teorema Central del Límite.

Aplicar el Teorema Central del Límite a las siguientes situaciones correspondiente a la suma de variables aleatorias independientes con igual distribución.

- a) El depósito de un avión se carga con 100 paquetes cuyos pesos son variables aleatorias independientes que están uniformemente distribuidas entre 5 y 50 kilos. ¿Cuál es la probabilidad de que el peso total de la carga exceda los 3000 kilos?
- b) Treinta y seis instrumentos electrónicos se usan de la siguiente manera: tan pronto falla el primero empieza a actuar el segundo, cuando éste falla empieza a actuar el tercero y así sucesivamente. Si la variable aleatoria asociada a la duración de cada instrumento, X_i , tiene una distribución exponencial de parámetro $0,1h^{-1}$, cada instrumento tiene un tiempo de operación independiente de los demás, y T es el tiempo total de operación de los 36 instrumentos, ¿cuál es la probabilidad de que T exceda las 370 hs? Justificar el planteo del cálculo.
- c) Sea $\{X_1, X_2, \dots, X_{50}\}$ un conjunto de variables aleatorias independientes que tienen cada una una distribución Poisson con parámetro $\lambda=0,3$, y sea $T_{50}=X_1+X_2+\dots+X_{50}$ su suma. Evaluar usando el Teorema Central del Límite $P(T_{50} \geq 18)$.
- d) **Tiempo de procesamiento de partes y cantidad de unidades.** Una máquina procesa partes, una a la vez. El tiempo de procesamiento de las diferentes partes son variables aleatorias independientes uniformemente distribuidas en el intervalo entre 1 y 5. Calcular la probabilidad de que el número de partes procesadas en un intervalo de 320 unidades de tiempo sea al menos 100.
- e) **Número de clientes atendidos en una caja.** Los tiempos de servicio para los clientes que llegan a una caja de supermercado son variables aleatorias independientes con un promedio de 1,5 minutos y una dispersión de 1 minuto.
- Calcular la probabilidad aproximada de que se pueda atender a 100 clientes en menos de 2 horas de tiempo total de servicio de esta caja.
 - Calcular el número de clientes n tal que la probabilidad de dar servicio a todos en menos de 2 horas sea aproximadamente 0,1.

Ejercicio 6. Aproximación normal de una variable aleatoria binomial.

Una variable aleatoria binomial S_n con parámetros n y p se puede ver como la suma de n variables aleatorias independientes con distribución Bernoulli, $\{X_1, X_2, \dots, X_n\}$, con parámetro común p :

$$S_n = X_1 + X_2 + \dots + X_n; E(X_i) = p; V(X_i) = p(1-p) \Rightarrow E(S_n) = np; V(S_n) = np(1-p).$$

Si n es lo suficientemente grande y el histograma de probabilidad binomial no está demasiado sesgado ($np \geq 10$ y $n(1-p) \geq 10$), entonces S_n tiene distribución aproximadamente normal con valor medio np y varianza $np(1-p)$.

A los efectos prácticos, se considera la conocida corrección de medio punto de modo que

$$P(S_n \leq a) = B(a; n, p) \cong P(S_n|_{NOR.} \leq a + 0,5) = P\left(Z \leq \frac{a + 0,5 - np}{\sqrt{np(1-p)}}\right) = \Phi\left(\frac{a + 0,5 - np}{\sqrt{np(1-p)}}\right),$$

$$P(a \leq S_n \leq b) \cong P(a - 0,5 \leq S_n|_{NOR.} \leq b + 0,5) = \Phi\left(\frac{b + 0,5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - 0,5 - np}{\sqrt{np(1-p)}}\right).$$

- a) Sea X una variable aleatoria binomial con parámetros $n=36$ y $p=0,5$. Calcular el valor de $P(X \leq 21)$ y comparar con él, el valor aproximado utilizando el teorema central del límite con y sin la corrección de medio punto.

- b) Para la misma variable del ítem anterior, comparar $P(X = 19)$ con el valor aproximado utilizando el teorema central del límite dado por

$$P(X = 19) \cong P(19 - 0,5 \leq X_{NOR} \leq 19 + 0,5) = \Phi\left(\frac{19,5 - 18}{3}\right) - \Phi\left(\frac{18,5 - 18}{3}\right).$$

Ejercicio 7 Promedio muestral y Teorema Central del Límite.

- a) Sea X_1, X_2, \dots, X_{40} una muestra aleatoria de mediciones de la proporción de impurezas en muestras de una aleación. Cada una de las variables aleatorias X_k tiene una densidad de probabilidad definida por

$$f(x) = \begin{cases} 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{en otro caso} \end{cases}.$$

Un comprador potencial rechaza una partida de mineral si la media de la proporción de impurezas de una muestra de tamaño 40 es superior a 0.8. Calcular la probabilidad de que la partida sea rechazada. Comentar claramente el planteo del problema.

- b) Una máquina automática de llenado de botellas tiene una varianza en las cantidades de llenado de $\sigma^2=1$ onza². Sin embargo, el promedio de las onzas de llenado μ depende de un ajuste diario y de operador a operador. Si en un día se llevan a cabo 36 observaciones de la cantidad (en onzas) del líquido de llenado y con el mismo ajuste de máquina, calcular:
- i. la probabilidad de que el promedio muestral quede dentro de 0,3 onzas de diferencia con respecto al promedio verdadero de la población, para ese ajuste.
 - ii. ¿Cuál debe ser el tamaño de la muestra n , para que \bar{X} quede a menos de 0,3 onzas de μ con una probabilidad de 0,95?
- c) En la fabricación de cierto tipo de cojinetes para motor se sabe que el diámetro promedio es de 5 cm con un desvío estándar de 0,005 cm. El proceso es controlado en forma periódica mediante la selección aleatoria de 64 cojinetes, midiendo sus correspondientes diámetros. El proceso se supone bajo control si la media muestral se encuentra entre dos límites especificados en el 95% de las veces que se extraen las muestras para realizar el control del proceso. Determinar los valores de esos límites si son simétricos respecto de la media poblacional.

Ejercicio 8. Se sabe que el monto de las ventas realizadas por una empresa tiene distribución normal con media \$343200 y desvío típico \$48152. De las 1815 ventas realizadas en el mes se saca una muestra de 16 facturas.

- a) ¿Cuál es la probabilidad de que la media muestral difiera de la media poblacional en más de \$20000?
- b) ¿Cuál es el valor de la media muestral que es superado con probabilidad 0,05?

Ejercicio 9. Se considera un evento A definido en el contexto de algún experimento aleatorio. Sea $p=P(A)$ la probabilidad de que el evento ocurra. Si se consideran n repeticiones independientes, se puede definir la **frecuencia empírica de ocurrencia de A** como

$$f_n = \frac{X_1 + X_2 + \dots + X_n}{n},$$

donde X_i es 1 si ocurre A en la repetición i -ésima, y es 0 en caso contrario, con $E(X_i)=p$. Así, $E(f_n)=p$ y $V(f_n)=p(1-p)/n$.

- a) Una fábrica produce determinados artículos de tal manera que el 5% resulta defectuoso. Se inspeccionan n artículos y se determina la frecuencia empírica f_n de defectuosos entre los seleccionados. ¿Cuál debe ser el tamaño n si se desea que f_n difiera de la verdadera proporción $p=0,05$ de defectuosos en menos de 0,01 (1 punto porcentual) con una probabilidad no menor a 0,95, esto es $P(|f_n - 0,05| < 0,01) \geq 0,95$?

- b) Si en el ítem anterior no se conociera el porcentaje p de defectuosos en lo producido, para evaluar el tamaño n de la muestra se utiliza que la dispersión será máxima cuando la función $p(1-p)$ tenga un máximo, esto es para $p=1/2$. A partir de esto, evaluar n para que $P(|f_n - p| < 0,01) \geq 0,95$.

Ejercicio 10. Diferencias de medias muestrales

- a) Se toman muestras independientes de tamaño 10 y 15 de una variable aleatoria distribuida normalmente con valor esperado 10 y varianza 3. ¿Cuál es la probabilidad de que el promedio de las dos muestras se diferencie (en valor absoluto) en más de 0,3?
- b) La vida útil en horas de un componente utilizado en la turbina de una aeronave es una variable aleatoria X con media $\mu_X=5000$ y desviación estándar $\sigma_X=40$, siguiendo una distribución normal. Se produce una mejora en el diseño del componente que modifica el tiempo de vida útil, variable aleatoria Y con distribución normal, pero de mayor valor medio, siendo ahora $\mu_Y=5050$ y disminuyen la desviación estándar a $\sigma_Y=30$. Suponiendo que el primer proceso utiliza una muestra aleatoria de $n_X = 16$ componentes y el mejorado una muestra aleatoria de $n_Y = 25$ elementos, ¿cuál es la probabilidad de que la diferencia de las dos medias muestrales $(\bar{X} - \bar{Y})$ sea al menos de 25 horas (para ser significativa)?

Ejercicio 11. Generación de números aleatorios.

- a) Generar 10 números al azar (muestra aleatoria de tamaño 10 provenientes de una distribución uniforme en el intervalo $[0; 1]$), sumarlos, restarle 5 a la suma y dividir por $\sqrt{10/12}$ (¿Un número mágico, tal vez?). Repetir esa generación 2000 veces, por ejemplo. Hacer un histograma de los 2000 números obtenidos y comparar con la distribución normal estándar.
- b) Generar las compras de una muestra de 9 clientes considerando que el pago efectuado por uno de ellos seleccionado al azar tiene la distribución de la variable aleatoria X definida en el ejercicio 1 de este trabajo práctico. Calcular el promedio muestral de los pagos, restarle $E(X)$ y dividirlo por $\sigma_X/3$. Repetir esa generación 2000 veces, por ejemplo. Hacer un histograma de los 2000 números obtenidos y comparar con la distribución normal estándar.
- c) Generar 50 muestras cada una de tamaño 10 de una variable aleatoria con distribución exponencial de parámetro $\alpha=2$. Para cada muestra calcular la media muestral obteniéndose así 50 realizaciones del estadístico media muestral para muestras de tamaño 10. Agrupar esos datos y analizar las frecuencias relativas. Repetir el procedimiento pero ahora con un tamaño de muestra igual a 500. Comentar los resultados.

Para este último ejercicio, y como introducción al trabajo práctico para el laboratorio de computación, se presenta aquí una introducción a la temática.

Generación de números aleatorios y su implementación computacional.

El desarrollo de los sistemas de cómputos ha convertido a los experimentos de simulación en técnicas útiles para el análisis de sistemas complejos. En la simulación de estos sistemas surge la necesidad de simular fenómenos aleatorios que le son característicos. Por ejemplo, si un negocio desea estudiar su política de servicios al cliente, puede simular el flujo de clientes como así también el tiempo necesario de atención a cada uno de ellos. Si una empresa desea examinar su política de mantenimiento de planta, puede simular el tiempo entre fallas de las componentes como así también el tiempo necesario para la reparación de cada una de ellas. Estos sucesos constituyen eventos aleatorios.

Para encarar este tipo de problemas se supone, en general, una distribución de probabilidad adecuada para cada fenómeno y se genera una secuencia de valores para la correspondiente variable aleatoria por computadora. La función de distribución uniforme en el intervalo $(0; 1)$ juega un papel importante en la generación de números aleatorios por computadora. Los sistemas de cómputos tienen en su estructura la capacidad de generar valores aleatorios que responden a este tipo de

distribución. A partir de ellos se puede generar números aleatorios con otras distribuciones de probabilidad específica, como se muestra en los siguientes ejemplos. Algunos *software* tiene incorporada esta posibilidad automáticamente, sin necesidad de conversiones a una función de distribución uniforme.

Ejemplo 1. Sea T una variable aleatoria continua que represente el tiempo entre dos fallas sucesivas de un dispositivo que sigue un proceso Poisson y se desea simular. La función de densidad

de probabilidad de T es exponencial con parámetro b , esto es $f(t) = \begin{cases} be^{-bt} & t \geq 0 \\ 0 & t < 0 \end{cases}$ con $b > 0$. La

función acumulada de probabilidad correspondiente es $F(t) = \begin{cases} 1 - e^{-bt} & t \geq 0 \\ 0 & t < 0 \end{cases}$. A partir de la función

RAND de calculadoras, o algún paquete estadístico para computadora, se puede obtener un número aleatorio entre 0 y 1 (que responde a una distribución uniforme entre dichos valores). Sea a el número sorteado. Este número se puede hacer corresponder con un valor de t , asociándolo con un valor de la función acumulada de distribución $F(t)$ y despejando en la forma $a = 1 - e^{-bt} \Rightarrow t = [-\ln(1 - a)]/b$. Así, por ejemplo, si $a = 0,5965$, $t = 90,7579$.

Ejemplo 2. Sea X una variable aleatoria discreta que represente las ventas diarias de un negocio. De los datos históricos se obtiene la siguiente distribución de probabilidad.

x	1	2	3	4
$p(x)$	0,20	0,40	0,30	0,10

Queda definida la función de distribución acumulada de la forma,

$$F(x) = \begin{cases} 0 & x < 1 \\ 0,20 & 1 \leq x < 2 \\ 0,60 & 2 \leq x < 3 \\ 0,90 & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

Con un número aleatorio a obtenido con la función *RAND* de la calculadora, o de algún paquete estadístico, se puede simular el número de ventas de un día en el negocio. Este número sorteado se puede hacer corresponder con un valor de la función de distribución acumulada de la variable ventas de un día X . Como puede observarse, quedan definidos los siguientes valores críticos $\{0; 0,20, 0,60; 0,90; 1\}$. Así, si $a \in [0, 0,20)$, entonces no se produce ninguna venta en el día, si $a \in [0,20, 0,60)$ entonces se produce 1 venta en el día, y así sucesivamente.

Ejemplo 3. Sea Z una variable aleatoria con distribución normal estandarizada, media 0 y desvío estándar 1. La función de distribución acumulada que le corresponde es

$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{1}{2}t^2\right] dt$. Se podría asociar esta expresión con un número aleatorio a obtenido

con la función *RAND* de la calculadora, como en los ejemplos anteriores. Sin embargo, la resolución de $F(x)=a$ no puede resolverse en forma cerrada para x (como en el ejemplo 1), esto es no existe forma analítica de despejar x en función de a . Una alternativa es recurrir a las tablas numéricas de $F(x)$. Otra alternativa es utilizar el hecho demostrable que si U_1 y U_2 son dos variables aleatorias independientes con distribución uniforme sobre el intervalo unitario, entonces

$$Z_1 = (-2 \ln U_1)^{1/2} \text{sen}(2\pi U_2) \text{ y } Z_2 = (-2 \ln U_1)^{1/2} \text{cos}(2\pi U_2)$$

son dos variables aleatorias normales estandarizadas e independientes. Cabe señalar que, con respecto a la distribución normal, muchos son los *software* que tienen incorporada la posibilidad de su generación en forma automatizada.

El equivalente a RAND en las calculadoras en el programa Excel es ALEATORIO(). El valor de la celda de esta función se refresca con la tecla F9. Para generar números aleatorios con otra distribución se puede explorar siguiendo la ruta Análisis de datos, Herramientas, Generación de números aleatorios.

Lectura y comprensión

Desigualdad de Tchebychev. Pafnuti Livovic Tchebychev, 1821-1894, fue uno de los primeros matemáticos del siglo XIX que intentó dar rigor al estudio de la Probabilidad. Dentro de los trabajos que realizó, uno de los más conocidos dentro del área es la desigualdad que lleva su nombre. Con ella se puede estimar cotas para probabilidades de eventos asociados con variables aleatorias a partir del conocimiento de su valor esperado y su varianza, sin conocer las características de la distribución de dicha variable incluyendo el no tener que tener en cuenta si es discreta o continua.

Esta desigualdad establece que si X es una variable aleatoria con valor esperado μ y varianza σ^2 , cualquiera sea la distribución, se verifica que

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}, \quad \forall c > 0.$$

Si $c = k\sigma$, la expresión es

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}, \quad \forall k > 0.$$

Si se toma una muestra aleatoria de X , $\{X_1, X_2, \dots, X_n\}$, se puede expresar la desigualdad para las variables aleatorias suma $T = X_1 + X_2 + \dots + X_n$ y promedio muestral $\bar{X} = (X_1 + X_2 + \dots + X_n)/n$, obteniéndose, respectivamente,

$$P(|T - n\mu| < k\sqrt{n}\sigma) \geq 1 - \frac{1}{k^2}, \quad \forall k > 0; \quad P\left(|\bar{X} - \mu| < k \frac{\sigma}{\sqrt{n}}\right) \geq 1 - \frac{1}{k^2}, \quad \forall k > 0.$$

Si se considera un evento A definido en el contexto de algún experimento aleatorio, con la probabilidad de que el evento ocurra indicada con $p = P(A)$, y se realizan n repeticiones independientes, se puede definir la frecuencia empírica de ocurrencia de A como

$$f_n = \frac{X_1 + X_2 + \dots + X_n}{n},$$

donde X_i es 1 si ocurre A en la repetición i -ésima, y es 0 en caso contrario con $E(X_i) = p$. La desigualdad de Tchebyshev se puede expresar

$$P(|f_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2}, \quad \varepsilon > 0.$$

Esta última expresión tiene una cota superior dada por

$$P(|f_n - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}, \quad \varepsilon > 0.$$

Ejemplos de aplicación.

Ejemplo 1. Un cañón efectúa disparos, siendo la probabilidad de dar en el blanco igual a 0,8 (se supone constante y los disparos independientes). Se efectúan 100 disparos y se desea estimar con la desigualdad de Tchebychev la probabilidad de que el número de blancos X satisfaga $70 \leq X \leq 90$.

$$X \sim \text{Bin}(n=100; p=0,8); E(X) = 80; V(X) = 16 \Rightarrow \sigma = 4$$

Usando la desigualdad de Tchebychev, en la forma $P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}, \forall k > 0$:

$$P(70 \leq X \leq 90) = P(|X - 80| \leq 2,5 \times 4) \geq 1 - \frac{1}{2,5^2} = 0,84$$

Se puede comparar este resultado con el cálculo usando la probabilidad acumulada de la distribución binomial usando las funciones del Excel:

$$= \text{DISTR.BINOM}(90,100,0.8,\text{VERDADERO}) - \text{DISTR.BINOM}(69,100,0.8,\text{VERDADERO})$$

$$P(70 \leq X \leq 90) = 0,99766644 - 0,00605934 = 0,9916071.$$

O bien, utilizar la aproximación de la distribución a normal, con el siguiente resultado:

$$P(70 \leq X \leq 90) \approx 0,99565.$$

Ejemplo 2. Si se comparan las $P(|X - \mu| < \sigma)$, $P(|X - \mu| < 2\sigma)$ y $P(|X - \mu| < 3\sigma)$ calculadas a partir la desigualdad de Tchebychev en forma general con las correspondientes a considerar una variable aleatoria X con distribución normal, se obtiene el siguiente cuadro:

	$P(X - \mu < \sigma)$	$P(X - \mu < 2\sigma)$	$P(X - \mu < 3\sigma)$
Des. Tchebychev	≥ 0	$\geq 0,75$	$\geq 0,8889$
Normal	0,6826	0,9544	0,9974

Ejemplo 3. Al sumar números, una computadora aproxima cada número al entero más próximo. Se supone que todos los errores de aproximación son variables aleatorias X_i independientes con distribución uniforme en el intervalo $(-0,5, 0,5)$. Si se estima con la expresión

$$P(|T - n\mu| \geq c) \leq \frac{n\sigma^2}{c^2}, \forall c > 0,$$

la probabilidad de que cuando se sumen 1500 números, el valor absoluto del error total T sea superior a 15, se obtiene que

$$P(|T| > 15) \leq 0,5555.$$

Ejemplo 4. Una fábrica produce determinados artículos de tal manera que el 5% resulta defectuoso. Se inspeccionan n artículos y se determina la frecuencia empírica f_n de defectuosos entre los seleccionados. Si se evalúa a partir de la desigualdad de Tchebychev cuál debe ser el tamaño n , para que f_n difiera de la verdadera proporción $p=0,05$ de defectuosos en menos de 0,01 (1 punto porcentual) con una probabilidad no menor a 0,95, esto es $P(|f_n - 0,05| < 0,01) \geq 0,95$, el resultado que se obtiene es $n \geq 9500$. En condiciones de total incertidumbre sobre el valor de p , $P(|f_n - p| < 0,01) \geq 0,95$, conduce a un tamaño muestral $n \geq 50000$. Ambos tamaños muestrales corresponden a sobre-estimaciones del valor buscado.