

Teorema del límite central

Escuela de Matemática

Semana 15

Giovanni
Sanabria Brenes

Luis Ernesto
Carrera Retana

Erick Chacón
Vargas

Mario Marín
Sánchez

Rebeca Solís
Ortega

Teorema del Límite central

El teorema del límite central es de suma importancia en diversas aplicaciones de la probabilidad. Si bien el hacer un estudio detallado de la teoría detrás del concepto de límite central está fuera del alcance de los objetivos de este curso, sí abordaremos algunos detalles que pueden ayudar al lector a obtener una mejor comprensión de los conceptos relacionados con este teorema y las aplicaciones que se derivan del mismo. Para hacer que la presentación sea más apegada a la realidad conceptual no se dará el concepto como una receta y se tratará de que el lector vaya llegando en forma gradual a los conceptos que se buscan.

1. Leyes de los Grandes Números

Dado un experimento con espacio muestral Ω , para un evento Y se ha indicado que si se hacen n repeticiones del experimento y se nota que en esas n repeticiones del experimento ocurren $Y(n)$ veces el evento Y , intuitivamente se tiene que:

$$P[Y] \approx \frac{Y(n)}{n}.$$

Sin embargo, como ya hemos apuntado antes esta definición deja abiertas una serie de preguntas. Por ejemplo si aceptamos definir la probabilidad como el valor límite de estos cocientes entonces la definición se complica. Primero que todo, qué garantiza que ese límite existe, segundo esta definición no es operacional en el sentido de que no es posible repetir infinitamente tal experimento. Estudiaremos *la ley de los grandes números* que nos ayudará a precisar un poco mejor el sentido de

$$P[Y] = \lim_{n \rightarrow \infty} \frac{Y(n)}{n}. \quad (1)$$

Simplificando un poco el problema, cada una de las repeticiones del experimento que se realicen en el contexto citado puede verse como un ensayo de Bernoulli donde el éxito coincide con la ocurrencia de Y . Así el número de éxitos X en los n ensayos del experimento es una variable aleatoria binomial en la cual la probabilidad de éxito es un valor desconocido p . Para esta variable sabemos que la media es np y la varianza es $np(1-p)$.

Si consideramos la variable aleatoria $Y = X/n$ es muy sencillo demostrar que la esperanza de Y es $np/n = p$ y que la varianza es $(np(1-p))/n^2 = p(1-p)/n$.

Aplicando la desigualdad de Chebyshev a Y con $t = \varepsilon$ obtenemos:

$$P \left[\left| \frac{X}{n} - p \right| \geq \varepsilon \right] \leq \frac{p(1-p)}{n\varepsilon^2}. \quad (2)$$

Es decir el límite (1) existe o dicho en palabras algo más simples dada cualquier precisión ε se puede encontrar un valor n de manera que el cociente éxitos entre el total de ensayos esté tan cerca del valor p desconocido como queramos.

En cierta forma esta última desigualdad da legitimidad al proceso estadístico pues garantiza que el proceso descrito en realidad converge al valor de la probabilidad del evento. Es más note que si a la desigualdad anterior se toma el límite cuando n tiende a infinito:

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{X}{n} - p \right| \geq \varepsilon \right] \leq \lim_{n \rightarrow \infty} \frac{p(1-p)}{n\varepsilon^2} = 0. \quad (3)$$

Lo que quiere decir que la probabilidad de que X/n y p sean diferentes tiende a 0 conforme n crece.

Por supuesto que no resuelve en forma simple el problema operacional de saber cuál debe ser el número de repeticiones del experimento necesarias para obtener aproximaciones precisas de la probabilidad buscada. Se puede utilizar la desigualdad de Chebyshev para obtener aproximaciones del valor de n pero el teorema del límite central, que abordaremos en la sección siguiente será de mayor utilidad en ese sentido.

Las conclusiones que se han obtenido hasta ahora se resumen en el siguiente teorema conocido como una forma débil de la *ley de los grandes números*.

Teorema 1.

Sea Y un evento y $Y(n)$ en número de ocurrencias del evento en n repeticiones del experimento. Entonces para todo $\varepsilon \geq 0$ se tiene

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{Y(n)}{n} - P[Y] \right| \geq \varepsilon \right] = 0. \quad (4)$$

Paralela a la forma débil de la ley de los grandes números existe una generalización que se llama la *Ley de los grandes Números* cuya justificación está fuera de los objetivos de este curso y se enuncia en el siguiente teorema:

Teorema 2.

Sean X_1, X_2, \dots, X_n variables aleatorias mutuamente independientes y todas siguiendo la misma distribución. Si existe la esperanza $\mu = E[X_k]$, entonces para todo $\varepsilon > 0$ se tiene

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mu \right| \geq \varepsilon \right] = 0. \quad (5)$$

Dicho en otras palabras la probabilidad de que el promedio $\bar{x} = S_n/n$ difiera de la esperanza en por lo menos un ε cualquiera, tiende a cero.

2. El Teorema del Límite Central

El último teorema de la sección previa es generalizado por otro teorema cuya importancia en aplicaciones de la probabilidad y estadística es mucho mayor. El teorema del límite central se enuncia seguidamente.

Teorema 3.

Sean X_1, X_2, \dots, X_n variables aleatorias mutuamente independientes y todas siguiendo la misma distribución. Si existen la esperanza $\mu = E[X_k]$ y la varianza $\sigma^2 = Var[X_k]$. Entonces para $S_n = X_1 + X_2 + \dots + X_n$ y $x < y$ se tiene

$$\lim_{n \rightarrow \infty} P \left[x \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq y \right] = \Phi(y) - \Phi(x). \quad (6)$$

Donde $\Phi(z)$ es la distribución normal estándar.

La importancia de este teorema es enorme, en especial porque no tiene ninguna condición especial sobre el tipo de distribución al que se aplica. Puede ser continua o discreta, no importa como sean, en promedio la suma de estas variables distribuyen como una normal con media $n\mu$ y varianza $n\sigma^2$. Este teorema también es válido para la variable aleatoria $\bar{X} = S_n/n$ para la que, si n se hace grande, distribuye como una normal de media μ_X y varianza σ_X^2/n

3. Estimadores

Definición 1.

Se dice que las variables aleatorias X_1, X_1, \dots, X_n forman una muestra aleatoria de tamaño n si son independientes dos a dos y todas siguen la misma distribución de probabilidad.

Definición 2.

Un estimador $\hat{\kappa}$ de un parámetro κ de una variable aleatoria X es una variable aleatoria, que puede depender de una muestra aleatoria X_1, X_1, \dots, X_n .

Los dos estimadores más usuales son el promedio usual llamado también media muestral y denotado por \bar{X} y la varianza muestral denotado por S^2

Estos estimadores son a su vez variables aleatorias,

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (7)$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (8)$$

La desviación estándar muestral S es la raíz de la varianza.

Como sus nombres lo indican, se tiene que \bar{X} es un estimador para la esperanza μ_X , S^2 lo es para la varianza $\text{Var}[X]$ y S para la desviación estándar σ_X .

El siguiente teorema, que en algunos textos se llama teorema del límite central, es sumamente útil pues permite resolver diversos ejercicios de manera bastante simple.

Teorema 4.

Sean X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n sobre una población que sigue una distribución dada por una variable aleatoria X con media μ y varianza σ^2 . Entonces se tiene:

1. $E[\bar{X}] = \mu$
2. $E[S^2] = \sigma^2$
3. $Var[\bar{X}] = \frac{\sigma^2}{n}$
4. Si n es suficientemente grande, entonces la variable

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (9)$$

sigue una distribución que se aproxima a una normal estándar.

5. Sea $S_n = X_1 + X_2 + \dots + X_n$. Si n es suficientemente grande, entonces la variable

$$Z = \frac{S_n - n\mu}{\sigma\sqrt{n}} \quad (10)$$

sigue una distribución que se aproxima a una normal estándar.

Para aplicar este teorema (llamado Teorema del Límite Central (TLC)) se tiene que si n es grande (empíricamente $n \leq 30$) entonces:

1. La distribución del promedio muestral \bar{X} se aproxima a la distribución $N(\mu, \frac{\sigma^2}{n})$.
2. La distribución de la suma S_n se aproxima a la distribución $N(n\mu, n\sigma^2)$.

Ejemplo 1.

Las consultas a un sistema tienen una duración cuya media es de 4 segundos y su desviación estándar es de 1.5 segundos. Si llegan 50 consultas en forma independiente, ¿cuál es la probabilidad de que las 50 tengan una duración promedio entre 3.5 y 3.8 segundos?

Solución: Sea X la duración de una consulta en segundos. Se sabe que $\mu = E(X) = 4$ y que $\sigma = \sqrt{Var(X)} = 1.5$.

Si aplicamos los resultados descritos hasta ahora el promedio de la muestra de las 50 consultas sigue una distribución que es aproximadamente normal con media $\mu = 4$ y desviación estándar $\sigma\sqrt{n} = 1.5/\sqrt{50} = 0.2121$. Es decir,

$$\bar{X} \sim (aprox)N(4, 0.2121^2)$$

Luego:

$$P[3.5 \leq \bar{X} \leq 3.8] = P\left[\frac{3.5 - 4}{0.2121} \leq Z \leq \frac{3.8 - 4}{0.2121}\right] = \Phi\left(\frac{3.8 - 4}{0.2121}\right) - \Phi\left(\frac{3.5 - 4}{0.2121}\right) = 0.1645.$$

Ejemplo 2.

Una sonda espacial cuenta con un juego de 10 computadores para controlar su estado. En todo momento se encuentra trabajando un único computador y estos trabajan en forma serial de manera que en el instante en que uno falle empieza a funcionar el siguiente, y así sucesivamente hasta utilizar los 10 computadores. La sonda está por pasar detrás de un planeta, por lo que se espera no tener comunicación con ella durante 4000 horas. Si Y es la variable el tiempo acumulado de funcionamiento de todos los computadores, determine la probabilidad de que el tiempo sea mayor a 4000 horas si:

1. cada computador opera correctamente 440 horas en promedio con una desviación estándar de 30 horas.
2. el promedio de funcionamiento de cada computador fuera de 410 horas y la desviación estándar de 30 horas.

Solución:

Sea X el tiempo que opera un computador.

1. Se sabe que $\mu = E(X) = 440$ y que $\sigma_X = 30$.

Entonces el tiempo acumulado de funcionamiento, $Y = S_{10}$, de todos los computadores sigue una distribución que se puede aproximar por una normal con media 4400 y desviación estándar $30\sqrt{10}$.

$$\begin{aligned} P[Y > 4000] &= P\left[Z > \frac{4000 - 4400}{30\sqrt{10}}\right] = 1 - \Phi\left(\frac{4000 - 4400}{30\sqrt{10}}\right) \\ &= 1 - \Phi(-4.21) \approx 1. \end{aligned}$$

2. Si el promedio de funcionamiento de cada computador fuera de 410 horas y la desviación estándar de 30 entonces la probabilidad pedida sería:

$$P[Y > 4000] = 1 - \Phi\left(\frac{4000 - 4100}{30\sqrt{10}}\right) = 1 - \Phi(-1.05409) = 1 - 0.14592 = 0.85408.$$

Ejemplo 3.

El tiempo de vida útil de una tablet marca XTEC sigue una distribución exponencial con una media de 3 años. Dada la variedad de marcas de tablet en el mercado, la organización Evalúa Tablet se ha dedicado a evaluar estos dispositivos y considera que una tablet es de buena calidad si tiene una vida útil mayor a 4 años.

1. Determine la probabilidad de que una tablet XTEC sea de buena calidad.
2. A partir del 2017 la empresa Evalúa Tablet otorgará un certificado de calidad a las empresas que fabrican tablet y que cumplan el siguiente test: al elegir 40 tabletas al azar, éstas deben ser, en promedio, de buena calidad. Determine la probabilidad de que XTEC obtenga el certificado de buena calidad.

Solución:

1. Se tiene que:

X : vida útil de tablet XTEC

$$X \sim \text{Exp}(1/3)$$

$$P(X > 4) = 1 - (1 - e^{-4/3}) \approx 0.263597$$

2. Note que

X : vida útil de tablet XTEC

$$E(X) = \frac{1}{\lambda} = 3, \quad \text{Var}(X) = \frac{1}{\lambda^2} = 9$$

$$TLC: \bar{X}_{\text{aprox}} \sim N\left(3, \frac{81}{40}\right)$$

$$P(\bar{X} > 4) = 1 - P\left(Z < \frac{4-3}{\sqrt{\frac{81}{40}}}\right) = 1 - \Phi(0.702728)$$

4. Aproximación Normal de Binomial

El teorema del límite central tiene una implicación adicional que también resulta sorprendente. Si S_n sigue una distribución binomial de parámetros n y p entonces si x y y son enteros no negativos tales que $x < y$, según el teorema del límite central se tiene que si n es suficientemente grande se cumple que:

$$P[x \leq S_n \leq y] \longrightarrow \Phi\left(\frac{y-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{x-np}{\sqrt{np(1-p)}}\right). \quad (11)$$

Este resultado se conoce como la aproximación normal de la binomial y dado que es aproximación continua de una distribución discreta deben tenerse algunos cuidados adicionales.

La mejor manera de utilizar este resultado puede obtenerse en la expresión:

$$\sum_{i=x}^y \binom{n}{k} p^k (1-p)^{n-k} \approx \Phi\left(\frac{y-np+\frac{1}{2}}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{x-np-\frac{1}{2}}{\sqrt{np(1-p)}}\right). \quad (12)$$

El valor $1/2$ que se agrega a cada lado se llama un factor de corrección de continuidad. La razón para agregar tal factor de corrección es que si uno usa una distribución normal, que es continua, para aproximar una binomial que es discreta, en cada extremo del intervalo la distribución discreta incluye la mitad de una barra que la distribución continua omite, por eso debe agregarse.

Empíricamente, si $X \sim B(n, p)$, para utilizar la aproximación por la Normal se debe tener que: $np \geq 5$ y $nq \geq 5$.

Ejemplo 4.

Se sabe que en una ciudad el 35 % de los habitantes tienen sobrepeso. Si se eligen 500 personas, ¿cuál es la probabilidad de que haya entre y 200 y 300 personas con sobrepeso?

Solución:

Sea X el número de habitantes con sobrepeso de 500. Note que

$$X \sim B(500, 0.35)$$

Por lo tanto, la solución de este problema se obtiene por la expresión

$$P(200 \leq X \leq 300) = \sum_{i=200}^{300} \binom{500}{k} (0.35)^k (0.65)^{500-k}$$

Sin embargo, como $np = 175 \geq 5$ y $nq = 325 \geq 5$, se tiene que por aproximación con la normal

$$X_{\text{aprox}} \sim N(175, 500(.35)(.65))$$

Así:

$$P(200 \leq X \leq 300) \approx \Phi\left(\frac{300 - 175 + \frac{1}{2}}{\sqrt{500(.35)(.65)}}\right) - \Phi\left(\frac{200 - 175 - \frac{1}{2}}{\sqrt{500(.35)(.65)}}\right).$$

Utilizando la fórmula para el cálculo de binomiales se obtiene que la parte derecha es 0.008864 mientras que la parte derecha, usando las tablas correspondientes, es 0.0108.

Existen varios criterios para asegurar la precisión de este tipo de aproximaciones. Los ejemplos abundan, por ejemplo unos autores afirman que si $np \geq 5$ y $n(1-p) \geq 5$ la aproximación es adecuada, mientras que otros presentan un resumen de diferentes condiciones para asegurar precisión. Al final de cuentas lo que si es válido es que valores de p muy cercanos a 0 o 1 hacen que las aproximaciones normales de binomiales no sean buenas.

Ejemplo 5.

El peso de una bolsa de tomate de cierta distribuidora alimenticia sigue una distribución normal con media de 1000 gramos y desviación estándar de 75 gramos. Debido a la sobre oferta de tomate, deciden hacer paquetes de 3 bolsas y vender cada paquete por 1000 colones. Un inspector decide revisar 1000 de tales paquetes. Si al menos 175 de ellos pesan menos de 2975 g, entonces castigará a la distribuidora con una multa. ¿Cuál es aproximadamente la probabilidad de que la distribuidora sea castigada?

Solución: Sea X el peso de 3 paquetes; así $X \sim N(3000, 625)$, por lo que

$$P(X < 2975) = \Phi\left(\frac{2975 - 3000}{25}\right) = \Phi(-1) = 0.15866.$$

Sea Y el número de paquetes que pesan menos de 2975 g, así $Y \sim B(1000, 0.15866)$, es decir, $Y_{\text{aprox.}} \sim N(158.66, 133.487)$, por lo que

$$P(Y \geq 175) = 1 - P(Y < 175) \approx 1 - \Phi\left(\frac{174.5 - 158.66}{\sqrt{133.487}}\right) \approx 1 - 0.92785 = 7.215\%.$$

Ejemplo 6.

El programa Quién Quiere ser Multimillonario, realizó un test a las personas interesadas en participar en la próxima temporada. Los resultados del test siguen una distribución normal con un promedio de 56 y una desviación estándar de 12, en escala de 1 a 100.

1. El 40 % de las personas que obtuvieron mejor resultado en el test, se considerarán candidatos elegibles para participar en la proxima temporada. Determine la nota mínima necesaria que de obtener una persona en el test, para ser considerada elegible.
2. El productor considera que la regla anterior no es atractiva pues para el ranking del programa se necesitan de todo tipo de personas. Así, considera que las personas elegibles se van tomando al azar, hasta obtener una probabilidad inferior al 5 % de que el resultado promedio, de las personas elegidas, sea menor a 50. ¿Cuántas personas considera que se deben elegir?

Solución.

1. Note que:

$$\begin{aligned}
 X &: \text{resultado en el test} \\
 X &\sim N(56, 12^2) \\
 c &: \text{nota mínima} \\
 P(X \geq c) = 40\% &\implies \phi\left(\frac{c-56}{12}\right) = 0.6 = \phi(0.25) \\
 \frac{c-56}{12} = 0.25 &\implies c = 59
 \end{aligned}$$

2. Se tiene que:

$$\begin{aligned}
 X &\sim N(56, 12^2) \implies \bar{X} \sim N(56, 12^2/n) \\
 P(\bar{X} \leq 50) &< 5\% \\
 \frac{50-56}{12/\sqrt{n}} &< -1.645 \\
 n &> 10.8241 \quad \text{Tome } n = 11
 \end{aligned}$$

Así la respuesta es $n = 11$. Note que como X es normal entonces \bar{X} es normal, sin importar el valor de n .

Si X no fuera normal, podemos resolver el ejercicio utilizando el TLC asumiendo que $n \geq 30$: \bar{X} se aproxima a $N(56, 12^2/n)$. Sin embargo, al obtener $n > 10.8241$, por el requisito del TLC, la respuesta sería $n=30$

5. Ejemplos propuestos

1. En una gran empresa el 60 % de las personas tiene problemas de tensión. ¿Cuál es la probabilidad de que en una muestra de 1000, 615 o más presenten este problema?

Solución:

Este problema es de tipo binomial, puede resolverse calculando en forma directa

$$1 - B(614; 1000, 0.6)$$

lo que conduce al valor 0.158528.

También podemos recurrir a la aproximación normal de binomial y la probabilidad solicitada es:

$$\begin{aligned} P[X \geq 615] &= \Phi\left(\frac{1000 - 600 + .5}{\sqrt{1000(0.6)(0.4)}}\right) - \Phi\left(\frac{615 - 600 - .5}{\sqrt{1000(0.6)(0.4)}}\right) \\ &= 1 - \Phi(.93597) = 0.174 \end{aligned}$$

2. El rendimiento de cierto cilindro de gas está normalmente distribuido con una media de 6 horas y una desviación estándar de 0.5 horas. Este gas se vende en paquetes de 5 cilindros y en cada paquete se utilizan los cinco cilindros en forma secuencial, es decir se empieza uno solamente si se ha terminado el anterior.

Determinar el tiempo máximo de duración de cada paquete de manera que éste sea excedido sólo por el 3 % de los paquetes.

Solución:

Como el tiempo de duración de cada cilindro es normal la distribución del tiempo $TP = T_1 + \dots + T_5$ de cada paquete también es normal con media 30 y desviación estándar $0.5\sqrt{5}$, lo que se solicita es un valor c tal que.

$$P[TP < c] = 0.97 = P\left[Z < \frac{c - 30}{0.5\sqrt{5}}\right] = 0.97$$

Utilizando las fórmulas correspondiente se obtiene

$$\frac{c - 30}{0.5\sqrt{5}} = 1.8807$$

de donde se concluye que $c = 31.977$, es decir solo un 3 % de los paquetes tienen una duración de más de 31.977 horas.

3. La duración de una batidora de un cierto fabricante es de 5 años, con una desviación estándar de un año. Si asumimos que las duraciones de estos mezcladores siguen aproximadamente una distribución normal, determine:
- La probabilidad de que en promedio este grupo dure entre 4.4 y 5.2 años.
 - El valor de \bar{x} a la derecha del cual caería el 15 % de las medias calculadas de una muestra aleatoria de tamaño 9.

Solución:

- a) Si se quiere la probabilidad de que en promedio este grupo dure entre 4.4 y 5.2 años se tiene

$$P[4.4 \leq \bar{X} \leq 5.2] = \Phi(0.60) - \Phi(-1.8) = 0.7257 - 0.0359 = 0.6898$$

- b) Si se toma una muestra aleatoria de 9 de estas batidoras entonces como la duración de un mezclador es de 5 años con una desviación de 1 año, la duración promedio sigue una distribución normal con una media de 5 años y con una desviación de $\frac{\sigma}{\sqrt{n}} = \frac{1}{3} = 0.3333$.

Luego el valor de \bar{x} a la derecha del cual caería el 15 % de las medias calculadas de las muestras aleatorias de tamaño 9 se obtiene del cálculo

$$P[\bar{X} \geq \bar{x}] = 0,15$$

o bien

$$P[\bar{X} \leq \bar{x}] = P\left[Z \leq \frac{\bar{x} - 5}{0.33333}\right] = 0,85$$

de la tabla y despejando se obtiene $\bar{x} = 5,35$, es decir si se compraran 9 batidoras un 15 % de éstas funcionaría por un período superior a 5.35 años.

4. Un médico atiende un paciente en un tiempo que es una variable aleatoria con media $\mu = 8$ minutos y desviación estándar 3 minutos. Si debe atender un total de 40 pacientes determine:
- La probabilidad de que atienda todos los pacientes en menos de 5 horas, asumiendo que los pacientes ingresan, en forma continua.
 - La probabilidad de que el tiempo promedio de atención sea superior a 7.5 minutos.

Solución:

Tenemos:

- La probabilidad de que atienda todos los pacientes en menos de 5 horas, asumiendo que los pacientes ingresan, en forma continua es:

$$P[T = T_1 + \dots + T_{40} \leq 300] = P\left[Z < \frac{300 - 320}{3\sqrt{40}}\right] = 0,1469$$

- La probabilidad de que el tiempo promedio de atención sea superior a 7.5 minutos se obtiene de:

$$P[\bar{X} > 7.5] = 1 - P\left[Z \leq \frac{7.5 - 8}{3/\sqrt{40}}\right] = 0.8531.$$

5. Una empresa vende 4 tipos de quesos, de acuerdo a la siguiente distribución de probabilidad:

Tipo de queso	Q_1	Q_2	Q_3	Q_4
Probabilidad	0.2	0.15	0.25	0.4
Precio en colones	2000	1500	1000	800

Si un día particular se venden 50 quesos, determine la probabilidad de que el ingreso promedio por esta venta sea superior a 1200 colones.

Solución: Se tiene:

$$P\left[\frac{x_1 + \dots + x_{50}}{n} > 1200\right] = 1 - P\left[\frac{x_1 + \dots + x_{50}}{n} < 1200\right]$$

y haciendo los respectivos cambios:

$$\Phi\left(\frac{1200 - 1195}{\frac{464.19}{\sqrt{50}}}\right) \approx 0.5319$$

Por lo que la probabilidad buscada es aproximadamente del 53 %.

6. Si un trabajo consiste de dos tareas consecutivas, el tiempo que tarda en hacerse la primer tarea sigue una distribución normal con media 20 minutos y varianza 4 minutos, mientras que el tiempo para realizar la segunda tarea también se distribuye en forma normal con media 40 minutos y desviación estándar 5 minutos. ¿Cuál es la probabilidad de que el trabajo se complete en menos de 62 minutos.?

Solución:

Como el tiempo para completar el trabajo es la suma de dos normales, es normal y su media es $\mu_T = 60$ y desviación estándar $\sigma_T = \sqrt{4 + 5^2} = \sqrt{29}$, luego

$$P[X < 62] = \Phi\left(\frac{62 - 60}{\sqrt{29}}\right) = \Phi(0.37) \approx 0.6443$$

7. En una pequeña industria de fábrica de cajetas de leche, por las estadísticas de años, se sabe que éstas pesan en promedio un kilo con una desviación estándar de 70 gramos. Si estas cajetas se empaacan en cajas de 50, ¿cuál es la probabilidad de que una de estas cajas pese más de 51 kilos?

Solución:

En este caso, como la muestra es grande no importa conocer la distribución inicial, de hecho $X = X_1 + \dots + X_{50}$ se puede aproximar usando una normal con media 50 y desviación estándar $\frac{7\sqrt{50}}{100}$ y la probabilidad pedida es

$$P[X > 51] = 1 - \Phi\left(\frac{51 - 50}{0.4949}\right) = 1 - \Phi(2.02) \approx 0.0217$$