Índice general

Capítulo 2

Distribuciones Muestrales

Para estudiar las características de la muestra existen varias posibilidades. Si la distribución muestral exacta es necesaria, entonces los métodos de transformación de variables aleatorias pueden ser utilizados. Aquí usaremos la técnica de la función característica ya que esta función siempre existe.

1. Función Característica

La Función Característica se que se define como la esperanza matemática de una variable aleatoria compleja, desempeña un papel importante en el cálculo de probabilidades, como instrumento analítico en las demostraciones de los teoremas de límites.

La Función Característica de una variable aleatoria es una función de variable real que toma valores complejos, que permite la aplicación de métodos analíticos, es decir, de análisis funcional en el estudio de la probabilidad. En este marco si se identifica la distribución de la variable aleatoria considerada con una medida positiva, la Función Característica se denomina transformada de Fourier de la medida correspondiente.

La Función Característica de una variable aleatoria, así como la función generadora de momentos, la función de distribución acumulada, la función de densidad o de masa definen la distribución de probabilidad de dicha variable aleatoria.

El uso de las propiedades de la Función Característica fue introducido en las probabilidades por Lyapunov en 1904 para la demostración del Teorema Central del Límite que hoy lleva su nombre. La versión definitiva de este teorema fue obtenida posteriormente por Lindeberg.

La Función Característica de una variable aleatoria se define como la esperanza matemática de la variable compleja e^{iXt} .

Definición 1.1. *Sea X una v.a.. La función característica de X es la función $\Phi : \mathbb{R} \longrightarrow \mathbb{C}$ definida por $\Phi(t) = \Phi_X(t) = E(e^{itX})$.*

donde se define

$$E\left(e^{itX}\right) = E[\cos(tX)] + iE[\operatorname{sen}(tX)], \quad t \in \mathbb{R}$$

Observación:

$$\begin{aligned}\Phi_X(t) &= \int_{-\infty}^{\infty} \cos(tX) dF_X(x) + i\int_{-\infty}^{\infty} \operatorname{sen}(tX) dF_X(x)\\ &= \int_{-\infty}^{\infty} e^{itX} dF_X(x), \quad t \in \mathbb{R}\end{aligned}$$

donde la última igualdad se deduce de la integral de Stieljes para el caso de integrandos complejos.

Propiedades de la Función característica:

1. $\Phi_X(0) = 1$

   Prueba. De la definición se sabe que,

   $$\Phi_X(0) = E[\cos(0X)] + iE[\text{sen}(0X)] = E[1] + iE[0] = 1$$

   $\square$

2. La función característica está limitada por 1.
   $|\Phi_X(t)| \leq 1, \forall t \in \mathbb{R}$

   Prueba. Recordemos que, si Y es una v.a. con una esperanza finita entonces $Var(Y) = E(Y^2) - \{E(Y)\}^2 \geq 0$, entonces se tiene que $\{E(Y)\}^2 \leq E(Y^2)$. Como $t \in \mathbb{R}$ y las variables $\cos(tX)$ y $\text{sen}(tX)$ son acotadas y en particular se tiene:

   $$\{E[\cos(tX)]\}^2 \leq E[\cos^2(tX)] \tag{1.1}$$
   $$\{E[\text{sen}(tX)]\}^2 \leq E[\text{sen}^2(tX)] \tag{1.2}$$

   al sumar la ecuación (1.1) con la ecuación (1.2) se obtiene que:

   $$\begin{aligned}\{E[\cos(tX)]\}^2 + \{E[\text{sen}(tX)]\}^2 &\leq E[\cos^2(tX)] + E[\text{sen}^2(tX)]\\ &= E[\cos^2(tX) + \text{sen}^2(tX)]\\ &= E[1] = 1\end{aligned}$$

   por consiguiente como

   $$|\Phi_X(t)| = \sqrt{\{E[\cos(tX)]\}^2 + \{E[\text{sen}(tX)]\}^2}$$

   de la desigualdad anterior se concluye que $|\Phi_X(t)| \leq 1$ $\square$

3. $\forall t \in \mathbb{R} : \overline{\Phi_X(t)} = \Phi_X(-t)$

Prueba. Sabemos que las funciones trigonométrica $\operatorname{sen}(\cdot)$ y $\cos(\cdot)$ son funciones impar y par respectivamente, desarrollando directamente resulta:

$$\begin{aligned}\forall t \in \mathbb{R} : \overline{\Phi_X(t)} &= E[\cos(tX)] - iE[\operatorname{sen}(tX)] \\ &= E[\cos(-tX)] + iE[\operatorname{sen}(-tX)] \\ &= \Phi_X(-t)\end{aligned}$$

$\square$

4. Si X y Y son variables aleatorias independientes, entonces $\Phi_{X+Y}(t) = \Phi_X(t)\Phi_Y(t)$; $\forall t \in \mathbb{R}$

   Prueba. Recordando las identidades trigonométricas:

$$\begin{aligned}\cos(\alpha+\beta) &= \cos(\alpha)\cos(\beta) - \operatorname{sen}(\alpha)\operatorname{sen}(\beta) \\ \operatorname{sen}(\alpha+\beta) &= \operatorname{sen}(\alpha)\cos(\beta) + \cos(\alpha)\operatorname{sen}(\beta)\end{aligned}$$

   se tiene que $(\forall t \in \mathbb{R})$:

$$\cos[t(X+Y)] = \cos(tX)\cos(tY) - \operatorname{sen}(tX)\operatorname{sen}(tY) \tag{1.3}$$

$$\operatorname{sen}[t(X+Y)] = \operatorname{sen}(tX)\cos(tY) + \cos(tX)\operatorname{sen}(tY) \tag{1.4}$$

   Por otro lado dado que $t \in \mathbb{R}$ como X y Y son variables aleatorias independientes, entonces por propiedad "hereditaria" de familias de variables aleatorias independientes, las funciones $\cos(tX)$ y $\cos(tY)$, $\operatorname{sen}(tX)$ y $\operatorname{sen}(tY)$ son funciones independientes y por consiguiente tienen esparanzas finitas. Así

$$E[\cos(tX)\cos(tY)] = E[\cos(tX)]E[\cos(tY)] \tag{1.5}$$

Análogamente se deduce que:

$$E[\text{sen}(tX)\,\text{sen}(tY)] = E[\text{sen}(tX)]E[\text{sen}(tY)] \tag{1.6}$$
$$E[\text{sen}(tX)\cos(tY)] = E[\text{sen}(tX)]E[\cos(tY)] \tag{1.7}$$
$$E[\cos(tX)\,\text{sen}(tY)] = E[\cos(tX)]E[\text{sen}(tY)] \tag{1.8}$$

Tomando esperanza en las ecuaciones (1.3), (1.4) y luego usando las igualdades (1.5),(1.6), (1.7) y (1.8) se botiene $(\forall t \in \mathbb{R})$:

$$\begin{aligned}
\Phi_{X+Y}(t) &= E\{\cos[t(X+Y)]\} + iE\{\text{sen}[t(X+Y)]\}\\
&= E[\cos(tX)]E[\cos(tY)] - E[\text{sen}(tX)]E[\text{sen}(tY)]\\
&\quad + iE[\text{sen}(tX)]E[\cos(tY)] + iE[\cos(tX)]E[\text{sen}(tY)]\\
&= E[\cos(tX)]\,\{E[\cos(tY)] + iE[\text{sen}(tY)]\}\\
&\quad + iE[\text{sen}(tX)]\,\{E[\cos(tY)] + iE[\text{sen}(tY)]\}\\
&= E[\cos(tX)]\Phi_Y(t) + iE[\text{sen}(tX)]\Phi_X(t)\\
&= \{E[\cos(tX)] + iE[\text{sen}(tX)]\}\,\Phi_Y(t)\\
&= \Phi_X(t)\Phi_Y(t)
\end{aligned}$$

$\therefore\ \Phi_{X+Y}(t) = \Phi_X(t)\Phi_Y(t)$

□

Observación 1.1. Inductivamente se puede demuestra que, si $X_1,\ldots,X_n$ son variables aleatorias independientes entonces

$$(\forall t \in \mathbb{R}) : \Phi_{X_1,\ldots,X_n}(t) = \Phi_{X_1}(t)\cdots\Phi_{X_n}(t)$$

Observación 1.2. si $X_1,\ldots,X_n$ son variables aleatorias independientes y además se distribuyen idénticamente (iid) entonces

$$(\forall t \in \mathbb{R}) : \Phi_{X_1,\ldots,X_n}(t) = \Phi_X(t)\cdots\Phi_X(t) = \prod_{j=1}^{n}\Phi_X(t) = [\Phi_X(t)]^n$$

5. Si $Y = aX + b$ entonces $\Phi_Y(t) = e^{itb}\Phi_X(t)$

   Prueba. Definamos las variables aleatorias $X_1 = aX$ y $X_2 = b$ como X_2 es una variable aleatoria constante es independiente de cualquier otra v.a., en particular de X_1. Aplicando la propiedad 4, obtenemos que,

$$\Phi_{X_1+X_2}(t) = \Phi_{X_1}(t)\Phi_{X_2}(t)$$

   Pero dado que $t \in \mathbb{R}$ es evidente que

$$\begin{aligned}\Phi_{X_1}(t) &= E[\cos(t.aX)] + iE[\operatorname{sen}(t.aX)]\\ &= E[\cos(at.X)] + iE[\operatorname{sen}(at.X)]\\ &= \Phi_X(at)\end{aligned}$$

   y además

$$\begin{aligned}\Phi_{X_2}(t) &= E[\cos(t.b)] + iE[\operatorname{sen}(t.b)]\\ &= E[\cos(tb)] + iE[\operatorname{sen}(tb)]\\ &= e^{itb}\end{aligned}$$

   luego

$$\Phi_{aX+b}(t) = e^{itb}\Phi_X(at)$$

   □

6. Si $E(|X|^n) < \infty$ entonces Φ_X posee n derivadas continuas y

$$\Phi_X^{(k)}(t) = \int (ix)^k e^{itx} dF_X(x), k = 1, 2, \ldots, n.$$

   En particular, $\Phi_X^{(k)}(0) = i^k E(X^k)$, de modo la función caracteristica es una especie de la función generadora de momentos.

Observación 1.3.

$$\Phi_X^{(k)}(0) = \left.\frac{d^k \Phi_X(t)}{dt^k}\right|_{t=0}$$

es la k-ésima derivada evaluada en el punto $t = 0$

Teorema 1.1. [Fórmula de la Inversión] Si $x - h$ y $x + h$ son cualesquiera puntos de continuidad de $F(x)$,el incremento sobre el intervalo entre ellos está dado por la fórmula:

$$F(x+h) - F(x-h) = \lim_{x \longrightarrow \infty} \frac{1}{\pi} \int_{-T}^{T} \frac{\operatorname{sen}(ht)}{t} e^{-itx} \Phi_X(t) dt$$

Teorema 1.2. [Teorema de la Unicidad de la Función Característica] A cada función característica le corresponde una función de distribución única que tiene aquella función característica. En símbolos,

$$F_Y = F_X \Leftrightarrow \Phi_X = \Phi_Y$$

Teorema 1.3. [Teorema de Helly-Bray] Sean F, F_1,F_2,... funciones de distribuciones. Si F_n converge débilmente para F, entonces

$$\int g(x) dF_n(x) \xrightarrow[n \to \infty]{} \int g(x) dF(x)$$

para toda función g continua y limitada $(g : \mathbb{R} \longrightarrow \mathbb{R})$

Observación 1.4. Si $X_n \xrightarrow{D} X$, entonces se deduce del teorema que, $\int g(x)dF_n(x) \to \int g(x)dF(x)$ para toda función g continua y limitada, es decir, $E[g(X_n)] \to E[g(X)]$. En particular, como las funciones $\cos(tx)$ y $\operatorname{sen}(tx)$ son funciones continuas y limitadas para t fijo, se tiene $E[\cos(tX_n)] \xrightarrow[n \to \infty]{} E[\cos(tX)]$ y $E[\operatorname{sen}(tX_n)] \xrightarrow[n \to \infty]{} E[\operatorname{sen}(tX)]$, de modo que $\Phi_{X_n}(t) \to \Phi_X(t)\ \forall t \in \mathbb{R}$.

Teorema 1.4. [Teorema de la Continuidad de Paul Lévy] Sean F_1,F_2,... funciones de distribuciones y Φ_1, Φ_2... respectivamente sus funciones características. Si Φ_n converge puntualmente para un límite Φ y si Φ es continua en el punto cero, entonces

(a) existe una función de distribución F tal que $F_n \to F$ débilmente y

(b) Φ es la función característica de F.

Observación 1.5. Cabe aclarar que los teoremas 1.3 y 1.4 implican que $X_n \to X \Leftrightarrow \Phi_{X_n} \to \Phi_X$. Pero el Teorema de la Continuidad es más fuerte que el de la suficiencia de esa proposición, porque afirma que el límite de una sucesión de funciones características también es una función característica, con tal de que sea continua en el punto cero.

Corolario 1.1. (a) Sean $X_1, X_2, \ldots$ variables aleatorias . Si $\Phi_{X_n}(t) \to \Phi(t)\, \forall t \in \mathbb{R}$ y si Φ es continua en el punto cero, entonces Φ es función característica de alguna variable, digamos $\Phi = \Phi_X$, y $X_n \xrightarrow{D} X$.

(b) Si $\Phi_{X_n}(t) \to e^{-\frac{t^2}{2}}\, \forall t$, entonces $X_n \xrightarrow{D} \mathscr{N}(0,1)$.

Notación. "$X_n \xrightarrow{D} \mathscr{N}(0,1)$" indica que X_n converge en distribución para una variable aleatoria X que posee distribución normal $\mathscr{N}(0,1)$. Pero no es necesario que X sea explícitamente definida, y se puede interpretar la expresión "$X_n \to \mathscr{N}(0,1)''$ como que indica la convergencia débil de F_{X_n} para Φ, la función distribución de la normal estándar, $\mathscr{N}(0,1)$. Es conveniente decir, en este caso, que X_n converge en distribución para la (distribución)normal-estándar. Vale una interpretación análoga para las expresiones "$X_n \xrightarrow{D}$ Poisson(λ)",

"$X_n \xrightarrow{D} \chi^2(1)$", etc..

Teorema 1.5. [Teorema Central del Límite para variables aleatorias independientes e idénticamente distribuidas] Sean $X_1, X_2, \ldots$independiente e idénticamente distribuidas, con media común μ y varianza común σ^2, donde $0 < \sigma^2 < \infty$. Sea $S_\mathbf{n} = \sum_{j=1}^{\mathbf{n}} X_j$. Entonces

$$\frac{S_\mathbf{n} - E(S_\mathbf{n})}{\sqrt{Var(S_\mathbf{n})}} = \frac{S_\mathbf{n} - \mathbf{n}\mu}{\sigma\sqrt{\mathbf{n}}} = \frac{\sqrt{\mathbf{n}}(\overline{X} - \mu)}{\sigma} \xrightarrow{D} \mathscr{N}(0,1).$$

Corolario 1.2. [Teorema Central del Limite de De Moivre y Laplace] Sea $S_\mathbf{n}$ el número de éxitos en $\mathbf{n}$ ensayos de Bernoulli, con de probabilidad éxito θ en cada ensayo, donde $0 < \theta < 1$. Entonces

$$\frac{S_\mathbf{n} - \mathbf{n}\theta}{\sqrt{\mathbf{n}\theta(1-\theta)}} \xrightarrow{D} \mathscr{N}(0,1)$$

2. Distribución Muestral de un Estadístico

Un estadístico $T = T(X_1, \ldots, X_n)$, como una función de observaciones muestrales que son variables aleatorias, es también una variable aleatoria. Su valor varía de muestra a muestra. El comportamiento de la variabilidad en sus valores está dado por su distribución de probabilidad. La distribución de probabilidad de T se llama *distribución muestral* de T.

La deducción teórica de la distribución muestral de T puede ser fácil o difícil (o tal vez imposible de hallar). A veces, lo más que se puede hacer es obtener una distribución empírica de T mediante simulación.

Dos de los estadísticos más importantes en inferencia estadística son la media, $\overline{X}$ y la varianza S^2 muestrales.

Sea $X_1, \ldots, X_n$ una m.a.s. de una variable aleatoria que proviene de una distribución poblacional para la cual la función característica existe, entonces la función característica de la media muestral está dada por

$$\Phi_{\overline{X}}(t) = E\left[e^{it(\frac{1}{n}\sum_{j=1}^{n} X_j)}\right] = E\left[\prod_{J=1}^{n} e^{i\frac{t}{n}X_j}\right] = \left[\Phi_X\left(\frac{t}{n}\right)\right]^n$$

Ejemplo 2.1. Sean $X_1, \ldots, X_n$ una m.as. de una población con función de distribución $\Gamma(\alpha, 1)$. Hallamos la fdp de $\overline{X}$ como sigue: sabemos que si $X \sim \Gamma(\alpha, 1)$, entonces su función característica es $\Phi_X(t) = (1 - it)^{-\alpha}$, luego la función característica de $\overline{X}$ resulta

$$\Phi_{\overline{X}}(t) = \left[\Phi_X\left(\frac{t}{n}\right)\right]^n = \left[\left(1 - \frac{it}{n}\right)^{-\alpha}\right]^n = \left(1 - \frac{it}{n}\right)^{-n\alpha},$$

así, por el Teorema de la unicidad de la función característica concluimos, que $\overline{X} \sim \Gamma(\mathbf{n}\alpha, \mathbf{n})$ ■

Si se requiere utilizar distribución asintótica, se puede recurrir ya sea al uso del Teorema Central del límite o corolario del Teorema de Slutsky o el Teorema de Cramér.

Consideremos por ejemplo, **n** observaciones independientes, $X_1, \ldots, X_{\mathbf{n}}$ de una variable aleatoria X. Sea k entero y positivo y suponga que $E(|X|^{2k}) < \infty$. Entonces, cada variable X_i^k tiene media $m_k = E(X_i^k)$ y varianza $\mathbf{V}ar(X_i^k) = m_{2k} - \{m_k\}^2$.

Por Teorema Central de Límite resulta que, si a_k es el k-ésimo momento muestral, entonces

$$\frac{\sum X_i^k - \mathbf{n}m_k}{\sqrt{\mathbf{n}(m_{2k} - m_k^2)}} = \sqrt{\mathbf{n}}\frac{a_k - m_k}{\sqrt{(m_{2k} - m_k^2)}} \xrightarrow{D} Z \qquad (2.1)$$

donde $Z \sim \mathcal{N}(0,1)$. Así, a_k se distribuye en forma normal asintótica, es decir $a_k \overset{asin}{\sim} \mathcal{N}(m_k, \frac{m_{2k} - m_k}{\mathbf{n}})$. En particular, la media muestral, $\overline{X} \overset{asin}{\sim} \mathcal{N}\left(\mu, \frac{\sigma^2}{\mathbf{n}}\right)$.

De una manera similar, pero con la ayuda de un argumento algo más difícil, es posible demostrar que el momento muestral central de orden k también se distribuye en forma asintótica normal. Consultar Cramer,[18]pág.365, para mayor detalle. De hecho se puede demostrar según ciertas condiciones, cualquier característica de la muestra basada en momentos (muestrales) es asintóticamente normal con parámetros que son idénticos con la característica de la población correspondiente; ver Cramer [18]pág.366-367.

Ejemplo 2.2. Sea X_1,X_2,... variables aleatorias normales iid, (μ, σ^2). También sea $\overline{X}$, la media muestral y S^2 la varianza muestral. Consideremos la variable aleatoria

$$T_n = \frac{\sqrt{\mathbf{n}}(\overline{X} - \mu)/\sigma}{S/\sigma} = \frac{U_{\mathbf{n}}}{V_{\mathbf{n}}}$$

donde $U_n \sim \mathcal{N}(0,1)$. Más adelante se determinará la distribución de T_n. Aquí usamos el corolario ítem (c) del Teorema Slutsky para demostrar que $T_n \xrightarrow{D} Z$, donde $Z \sim \mathcal{N}(0,1)$. Se sabe que $S^2 \xrightarrow{P} \sigma^2$, de manera que $\frac{S}{\sigma} \xrightarrow{P} 1$. Luego podemos concluir que $T_n \xrightarrow{D} Z$.

El siguiente resultado de la distribución asintótica del r-ésimo estadístico de orden, $1 \leq r \leq n$, en el muestreo de una población con función de distribución F absolutamente continua con función de densidad f.

$$\left\{\frac{n}{p(1-p)}\right\}^{1/2} f(\zeta_p)(X_{(r)}-\zeta_p) \xrightarrow{D} Z, \text{ cunado } n \longrightarrow \infty$$

de modo que r/n permanece fijo, $r/n = p$, donde $Z \sim \mathcal{N}(0,1)$, y ζ_p es la única solución de $F(\zeta_p) = p$ (esto, es ζ_p es el único cuantil poblacional de orden p) ■

Observación 2.1. El cuantil muestral de orden p, Z_p, es asintóticamente normal,

$$Z_p \overset{asin}{\sim} \mathcal{N}\left(\zeta_p, \frac{1}{[f(\zeta_p)]^2}\frac{p(1-p)}{\mathbf{n}}\right),$$

donde ζ_p es el cuantil poblacional correspondiente, y f es la función de densidad de probabilidad de la función de distribución poblacional. Resulta que $Z_p \xrightarrow{p} \zeta_p$.

3. Muestreo de un Población Normal

Supongamos que $X_1, \ldots, X_{\mathbf{n}}$ es una m.a.s. de una población normal con media μ y varianza σ^2 , esto es $X \sim \mathcal{N}(\mu, \sigma^2)$. El objetivo es obtener las distribuciones de algunos estadísticos que son funciones de $\overline{X}$ y S^2.

Lema 3.1. Si una función característica bivariada se factoriza de modo que un factor es una función característica marginal:

$$\Phi_{U;V}(s,t) = \mathit{\Phi}_U(s)\Psi_V(t)$$

entonces el otro factor es la otra función característica marginal. Esto es $\Psi_V(t)$ es la función característica de V, y U y V son independientes. (Cualquiera o ambos U y V pueden ser vectores)

Teorema 3.1. (Teorema fundamental de muestreo de una población normal) Sea $X_1, \ldots, X_{\mathbf{n}}, \mathbf{n} \geq 2$, una m.a.s. de una distribución normal $\mathscr{N}(\mu, \sigma^2)$. Entonces

(a) La media muestral, $\overline{X}$ se distribuye en forma normal con media μ y varianza $\frac{\sigma^2}{\mathbf{n}}$, esto es, $\overline{X} \sim \mathscr{N}\left(\mu, \frac{\sigma^2}{\mathbf{n}}.\right)$

(b) $Z = \dfrac{\sqrt{\mathbf{n}}(\overline{X} - \mu)}{\sigma} \sim \mathscr{N}(0, 1)$.

(c) $\overline{X}$ y S^2 son independientes.

(d) $\dfrac{(\mathbf{n}-1)S^2}{\sigma^2}$ tiene distribución Ji-cuadrado con (n-1) grados de libertad; esto es, $\dfrac{(\mathbf{n}-1)S^2}{\sigma^2} \sim \chi^2(\mathbf{n}-1)$

(e) $t = \dfrac{\sqrt{\mathbf{n}}(\overline{X} - \mu)}{S} \sim t(\mathbf{n}-1)$

Prueba. (1) **Prueba del item (a):** Utilizando la función característica se sabe que , si $X \sim \mathscr{N}(\mu, \sigma^2)$ entonces su función característica es $\Phi_X(t) = e^{\mu it - \frac{\sigma^2 t^2}{2}}$. Luego la función característica de la media muestral es

$$\Phi_{\overline{X}}(t) = \left[e^{\mu i \dfrac{t}{\mathbf{n}} - \dfrac{\sigma^2 t^2}{2\mathbf{n}^2}} \right]^{\mathbf{n}} = e^{it\mu - \frac{\sigma^2}{\mathbf{n}} \frac{t^2}{2}}$$

y por el teorema de la unicidad de la función característica concluimos que $\overline{X} \sim \mathscr{N}(\mu, \frac{\sigma^2}{\mathbf{n}})$

(2) **Prueba del item (b):** Se sabe que si $X \sim \mathscr{N}(0, 1)$ entonces su función característica es $\Phi_X(t) = e^{-\frac{t^2}{2}}$ (prueba ejercicio).

Utilizando este resultado tenemos:

$$Z = \frac{\sqrt{\mathbf{n}}(\overline{X} - \mu)}{\sigma} = \frac{\sqrt{\mathbf{n}}}{\sigma}\overline{X} - \sqrt{\mathbf{n}}\frac{\mu}{\sigma}$$

y utilizando la propiedad 5 de la función característica tenemos

$$\Phi_Z(t) = \Phi_{\frac{\sqrt{\mathbf{n}}}{\sigma}\overline{X} - \sqrt{\mathbf{n}}\frac{\mu}{\sigma}}(t) = e^{-it\sqrt{\mathbf{n}}\frac{\mu}{\sigma}}\Phi_{\overline{X}}\left(\frac{\sqrt{n}}{\sigma}t\right)$$

y como

$$\Phi_{\overline{X}}\left(\frac{\sqrt{\mathbf{n}}}{\sigma}t\right) = e^{i\frac{\sqrt{\mathbf{n}}}{\sigma}t\mu - \frac{\sigma^2}{\mathbf{n}}\frac{n}{\sigma^2}\frac{t^2}{2}} = e^{i\frac{\sqrt{\mathbf{n}}}{\sigma}t\mu - \frac{t^2}{2}}$$

obtenemos

$$\Phi_Z(t) = e^{-i\frac{\sqrt{\mathbf{n}}}{\sigma}t\mu}e^{i\frac{\sqrt{\mathbf{n}}}{\sigma}t\mu - \frac{t^2}{2}} = e^{-\frac{t^2}{2}}$$

y otra vez, utilizando la propiedad de la unicidad de la función característica se concluye que

$$Z = \frac{\sqrt{\mathbf{n}}(\overline{X} - \mu)}{\sigma} \sim \mathscr{N}(0, 1).$$

(3) **Prueba de item (c):** Primero demostremos que $\overline{X}$ es independiente de $(X_j - \overline{X})$ para todo j. Para esto calculamos la función característica conjunta de $\overline{X}$ y del vector $(X_1 - \overline{X}), (X_2 - \overline{X}), \ldots, (X_{\mathbf{n}} - \overline{X})$; tenemos

$$\Phi_{\overline{X}, X_1 - \overline{X}, \ldots, X_{\mathbf{n}} - \overline{X}}(t, t_1, t_2, \ldots, t_{\mathbf{n}}) = \mathbf{E}\left(e^{it\overline{X} + i\sum_{j=1}^{\mathbf{n}} t_j(X_j - \overline{X})}\right) \tag{3.1}$$

Para continuar, desarrollemos el exponente de e en la ecuación (3.1), así tenemos,

$$\begin{aligned}
it\overline{X} + i\sum_{j=1}^{\mathbf{n}} t_j(X_j - \overline{X}) &= i\frac{t}{\mathbf{n}}\sum_{J=1}^{\mathbf{n}} X_j + i\sum_{J=1}^{\mathbf{n}} t_jX_j - i\sum_{j=1}^{\mathbf{n}} t_j\frac{1}{\mathbf{n}}\sum_{J=1}^{\mathbf{n}} X_j \\
&= i\frac{t}{\mathbf{n}}\sum_{J=1}^{\mathbf{n}} X_j + i\sum_{J=1}^{\mathbf{n}} t_jX_j - \bar{t}i\left(\frac{1}{\mathbf{n}}\right)\sum_{J=1}^{\mathbf{n}} X_j \\
&= i\frac{t}{\mathbf{n}}\sum_{J=1}^{\mathbf{n}} X_j + i\sum_{J=1}^{\mathbf{n}} X_j(t_j - \bar{t}) \\
&= i\sum_{J=1}^{\mathbf{n}} X_j\left\{\frac{t}{\mathbf{n}} + (t_j - \bar{t})\right\}
\end{aligned}$$

luego la ecuación (3.1) queda como sigue

$$\begin{aligned}
\Phi_{\overline{X},X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}}(t, t_1, t_2, \ldots, t_{\mathbf{n}}) &= \mathbf{E}\left[e^{i\sum_{J=1}^{\mathbf{n}} X_j\left\{\frac{t}{\mathbf{n}}+(t_j-\bar{t})\right\}}\right] \\
&= \mathbf{E}\left[\prod_{j=1}^{\mathbf{n}} e^{iX_j\left\{\frac{t}{\mathbf{n}}+(t_j-\bar{t})\right\}}\right]
\end{aligned}$$

como las variables X_j son iid (variables muestrales) resulta

$$\begin{aligned}
\Phi_{\overline{X},X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}}(t, t_1, t_2, \ldots, t_{\mathbf{n}}) &= \prod_{j=1}^{\mathbf{n}} E\left[e^{iX\left\{\frac{t}{\mathbf{n}}+(t_j-\bar{t})\right\}}\right] \\
&= \prod_{j=1}^{\mathbf{n}} \Phi_X\left(\frac{t}{\mathbf{n}} + (t_j - \bar{t})\right)
\end{aligned}$$

y como $X \sim \mathscr{N}(\mu, \sigma^2)$ tenemos

$$\begin{aligned}
\Phi_{\overline{X},X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}}(t, t_1, t_2, \ldots, t_{\mathbf{n}}) &= \prod_{j=1}^{\mathbf{n}} e^{i\left(\frac{t}{\mathbf{n}}+(t_j-\bar{t})\right)\mu-\frac{\sigma^2}{2}\left(\frac{t}{\mathbf{n}}+(t_j-\bar{t})\right)^2} \\
&= e^{\sum_{j=1}^{\mathbf{n}}\left[i\left(\frac{t}{\mathbf{n}}+(t_j-\bar{t})\right)\mu-\frac{\sigma^2}{2}\left(\frac{t}{\mathbf{n}}+(t_j-\bar{t})\right)^2\right]} \\
\Phi_{\overline{X},X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}}(t, t_1, t_2, \ldots, t_{\mathbf{n}}) &= e^{it\mu-\frac{\sigma^2}{\mathbf{n}}\frac{t^2}{2}-\frac{\sigma^2}{2}\sum_{j=1}^{\mathbf{n}}\left(t_j-\bar{t}\right)^2} \\
\Phi_{\overline{X},X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}}(t, t_1, t_2, \ldots, t_{\mathbf{n}}) &= \underbrace{e^{it\mu-\frac{\sigma^2}{\mathbf{n}}\frac{t^2}{2}}}_{\Phi_{\overline{X}}(t)}\ \underbrace{e^{-\frac{\sigma^2}{2}\sum_{j=1}^{\mathbf{n}}\left(t_j-\bar{t}\right)^2}}_{\Psi_{X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}}(t_1,t_2,\ldots,t_{\mathbf{n}})}
\end{aligned}$$

es decir, la función característica conjunta se factoriza

$$\Phi_{\overline{X},X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}}(t,t_1,t_2,\ldots,t_{\mathbf{n}}) = \mathit{\Phi}_{\overline{X}}(t).\Psi_{X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}}(t_1,t_2,\ldots,t_{\mathbf{n}})$$

y como el primer factor de lado derecho de la última igualdad es la función característica de la media muestral, $\overline{X}$, que proviene de la $\mathscr{N}(\mu,\sigma^2)$, según el lema el otro factor es función característica del vector $X_1-\overline{X},\ldots,X_n-\overline{X}$ y $\overline{X}$ y por el mismo lema $X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}$, entonces, la función característica conjunta de $\overline{X}$ y del vector $X_1-\overline{X},\ldots,X_n-\overline{X}$ y $\overline{X}$ y por el mismo lema $X_1-\overline{X},\ldots,X_{\mathbf{n}}-\overline{X}$, se factoriza, luego por el lema 3.1 ellos son independientes.

Como la varianza muestral, S^2 es función del vector $(X_1-\overline{X},\ldots,X_n-\overline{X})$ que es independientes de $\overline{X}$y por la propiedad "hereditaria de variables independientes"(Proposición 2.8 barry r. james, pag. 72) se concluye que $\overline{X}$ y S^2 son independientes.

(4) **Prueba del item (d):** En la prueba utilizamos la independencia de $\overline{X}$ y de S^2. Tenga en cuenta que,

$$\mathbf{n}\left(\frac{\overline{X}-\mu}{\sigma}\right)^2+\frac{(\mathbf{n}-1)S^2}{\sigma^2}=\sum_{i=1}^{\mathbf{n}}\left(\frac{X_i-\mu}{\sigma}\right)^2$$

De las propiedades de la distribuciones normales, se sabe que $\mathbf{n}\left(\frac{\overline{X}-\mu}{\sigma}\right)^2$ tiene distribución $\chi^2(1)$ cuya función característica es $(1-2it)^{-1/2}$ y $\sum_{i=1}^{\mathbf{n}}\left(\frac{X_i-\mu}{\sigma}\right)^2$ tiene distribución $chi^2(\mathbf{n})$ cuya función característica es $(1-2it)^{-\mathbf{n}/2}$. Por independencia de $\overline{X}$ yS^2, la función característica de $(\mathbf{n}-1)S^2/\sigma^2$

es

$$\Phi_{(\mathbf{n}-1)S^2/\sigma^2}(t) = \frac{(1-2it)^{-\mathbf{n}/2}}{(1-2it)^{-1/2}} = (1-2it)^{-\dfrac{(\mathbf{n}-1)}{2}}$$

Esta es la función característica de la distribución $\chi^2(\mathbf{n}-1)$. por lo tanto, el resultado está completo.

(5) **Prueba del item (e):** como $\dfrac{\sqrt{\mathbf{n}}(\overline{X}-\mu)}{\sigma} \sim \mathscr{N}(0,1)$, y $\dfrac{(\mathbf{n}-1)S^2}{\sigma^2} \sim \chi^2(\mathbf{n}-1)$ y $\overline{X}$ y S^2son independientes , entonces por definición resulta

$$\frac{\dfrac{\sqrt{\mathbf{n}}(\overline{X}-\mu)}{\sigma}}{\sqrt{\left[\dfrac{(\mathbf{n}-1)S^2}{\sigma^2}\right]/(\mathbf{n}-1)}} = \frac{\sqrt{\mathbf{n}(\overline{X}-\mu)}}{S} \sim t(\mathbf{n}-1)$$

□

4. Distribuciones Exactas:Chi-cuadrado, t, y F

En esta sección se investiga las distribuciones que surgen en el muestreo de una población normal. Consideremos una muestra aleatoria simple $X_1, \ldots, X_n$ extraída de una población normal $\mathscr{N}(\mu, \sigma^2)$. Entonces sabemos que $\overline{X} \sim \mathscr{N}(\mu, \sigma^2/n)$, probaremos que $\left\{\sqrt{n}(\overline{X}-\mu)/\sigma\right\}^2 \sim \chi^2(1)$.

La primera distribución de interés es la *distribución chi-cuadrado*, que tambien se puede definir como un caso especial de la distribución Gamma.

Definición 4.1. Una variable aleatoria X tiene una distribución chi-cuadrado (distribución-χ^2) con n grados de libertad si su función de densidad de probabilidad está dado por

$$f(x;\mathbf{n}) = \begin{cases} \frac{x^{\frac{\mathbf{n}}{2}-1}e^{-\frac{1}{2}x}}{2^{\frac{\mathbf{n}}{2}}\mathbf{\Gamma}(\frac{\mathbf{n}}{2})}, & \text{si } x > 0, \\ 0, & \text{si } x \leq 0 \end{cases}$$

Notación: $X \sim \chi^2(\mathbf{n})$ que se lee: la variable aleatoria X se distribuye como una v.a. Chi-cuadrado con n grados de libertad.

Nota: Si $X \sim \chi^2(\mathbf{n})$, entonces

Media: $E(X) = \mathbf{n}$

Varianza: $Var(X) = 2\mathbf{n}$

Asimetría: $\beta_1 = 2\sqrt{2/\mathbf{n}}$

Curtosis: $\beta_2 = 3 + \frac{12}{\mathbf{n}}$

Función Característica (F.C.) $\Phi_X(t) = (1-2it)^{-\frac{\mathbf{n}}{2}}$

Teorema 4.1. Sean $X_1, \ldots, X_{\mathbf{n}}$ variables aleatorias independientes e idénticamente distribuidas y sea $S_{\mathbf{n}} = \sum_{j=1}^{\mathbf{n}} X_j$. Entonces

(a) $S_{\mathbf{n}} \sim \chi^2(\mathbf{n}) \iff X_1 \sim \chi^2(1)$ y

(b) $X_1 \sim \mathscr{N}(0,1) \Longrightarrow \sum_{j=1}^{\mathbf{n}} X_j \sim \chi^2(\mathbf{n})$

Prueba. (1) **Prueba del item (a):**

$\Longrightarrow$ Supongamos que $S_{\mathbf{n}} \sim \chi^2(\mathbf{n})$, entonces la función característica de $S_{\mathbf{n}}$ es $\Phi_{S_{\mathbf{n}}}(t) = (1-2it)^{-\frac{\mathbf{n}}{2}} = \prod_{J=1}^{\mathbf{n}}(1-2it)^{-\frac{1}{2}}$, entonces cada factor es una función característica de una variable X_j, esto es $\Phi_{X_j}(t) = (1-2it)^{-\frac{1}{2}}$ para $j = 1, 2, \ldots, \mathbf{n}$ y así por el teorema de la unicidad de la función característica $X_1 \sim \chi^2(1)$

$\Longleftarrow$ si $X_1 \sim \chi^2(1)$, entonces la función característica de X_1 es $\Phi_{X_1}(t) = (1-2it)^{-\frac{1}{2}}$ y como las variables $X_1, \ldots, X_{\mathbf{n}}$ son iid se tiene que,

$$\Phi_{S_{\mathbf{n}}}(t) = \Phi_{\sum_{j=1}^{\mathbf{n}} X_j}(t) = \prod_{j=1}^{\mathbf{n}} (1-2it)^{-\frac{1}{2}} = (1-2it)^{-\frac{\mathbf{n}}{2}}$$

y nuevamente por el teorema de la unicidad de la función característica se concluye que $S_{\mathbf{n}} \sim \chi^2(\mathbf{n})$.

(2) **Prueba del (b):** Si $X_1 \sim \mathcal{N}(0,1) \Longrightarrow X_1^2 \sim \chi^2(1)$, y como las variables aleatorias $X_1, X_2, \ldots, X_{\mathbf{n}}$ son independientes e idénticamente distribuidas, por el ítem (a) se concluye que $\sum_{j=1}^{\mathbf{n}} X_j^2 \sim \chi^2(\mathbf{n})$.

$\square$

Corolario 4.1. Si $X_1, \ldots, X_{\mathbf{n}}$ es una muestra aleatoria simple de tamaño $\mathbf{n}$ de una población con distribución normal,$\mathcal{N}(\mu, \sigma^2)$. Entonces la variable aleatoria

$$Z = \sum_{J=1}^{\mathbf{n}} \left(\frac{X_j - \mu}{\sigma} \right)^2 = \frac{\sum_{J=1}^{\mathbf{n}} (X_j - \mu)^2}{\sigma^2} \sim \chi^2(\mathbf{n})$$

es decir tiene una distribución chi-cuadrado con $\mathbf{n}$ grados de libertad.

La distribución $\chi^2(\mathbf{n})$ está tabulado para valores de $\mathbf{n} = 1, 2, 3, \ldots$. Las tablas por lo general están hechas por hasta $\mathbf{n} = 30$, puesto que para $\mathbf{n} > 30$ se puede usar la aproximación normal.

Teorema 4.2. [Fisher] Si $X \sim \chi^2(\mathbf{n})$, entonces

$$\lim_{n \longrightarrow \infty} P\left(\sqrt{2X} - \sqrt{2\mathbf{n} - 1} \leq z\right) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt. \tag{4.1}$$

Prueba. Puesto que, X es la suma de $\mathbf{n}$ variables aleatorias independientes e idénticamente distribuidas como $\chi^2(1)$, aplicamos el Teorema Central de Límite para ver que

$$Z_{\mathbf{n}} = \frac{X - \mathbf{n}}{\sqrt{2\mathbf{n}}}$$

es normal asintótico, esto es,

$$\lim_{\mathbf{n} \longrightarrow \infty} P\left(Z_{\mathbf{n}} \leq z\right) = \lim_{\mathbf{n} \longrightarrow \infty} P\left(\frac{X - \mathbf{n}}{\sqrt{2\mathbf{n}}} \leq z\right) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

Luego,

$$\begin{aligned}
\lim_{\mathbf{n} \to \infty} P(\sqrt{2X^2} - \sqrt{2\mathbf{n} - 1} \leq z) &= \lim_{\mathbf{n} \to \infty} P(2X^2 \leq (z + \sqrt{2\mathbf{n} - 1})^2) \\
&= \lim_{\mathbf{n} \to \infty} P\left(X^2 \leq \frac{z^2 + 2\mathbf{n} - 1}{2} + z\sqrt{2\mathbf{n} - 1}\right) \\
&= \lim_{\mathbf{n} \to \infty} P\left(X^2 \leq \mathbf{n} + z\sqrt{2n}\right) \\
&= \lim_{\mathbf{n} \to \infty} P\left(\frac{X^2 - \mathbf{n}}{\sqrt{2\mathbf{n}}} \leq z\right) \\
&= \int_{-\infty}^{z} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt
\end{aligned}$$

$\square$

Observación 4.1. Resulta que para $z > 0$, si $X \sim \chi^2(\mathbf{n})$ se cumple

$$\begin{aligned} P(X \leq z) &= P(2X \leq 2z) = P(\sqrt{2X} \leq \sqrt{2z}) \\ &= P(\sqrt{2X} - \sqrt{2\mathbf{n}-1} \leq \sqrt{2z} - \sqrt{2\mathbf{n}-1}) \\ &= \int_{-\infty}^{\sqrt{2z}-\sqrt{2\mathbf{n}-1}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \end{aligned} \tag{4.2}$$

Se escribirá $\chi^2()\mathbf{n}, \alpha$ para el punto porcentual α superior de la distribución $\chi^2(\mathbf{n})$, es decir

$$P(\chi^2(\mathbf{n}) > \chi^2(\mathbf{n}, \alpha)) = \alpha.$$

Ejemplo 4.1. Sea $\mathbf{n} = 25$ y $X \sim \chi^2(25)$, entonces a partir de una tabla estadística chi-cuadrado o usando el aplicativo se tiene

$$P(X \leq 34{,}4) = 0{,}90.$$

aproximemos esta probabilidad por tablas normales usando el teorema de Fisher y lo comparamos con la aproximación del TCL. De la ecuación (4.2) tenemos

$$\begin{aligned} P(X \leq 34{,}4) &\approx P(Z \leq \sqrt{2 \times 34{,}382} - \sqrt{2 \times 25 - 1}) \\ P(X \leq 34{,}4) &\approx P(Z \leq \sqrt{68{,}764} - \sqrt{50-1}) \\ &= P(Z \leq 8{,}287 - 7) \\ &= P(Z \leq 1{,}29) = 0{,}9015; \text{ done } Z \sim \mathcal{N}(0,1). \end{aligned}$$

Para usar el TCL. vemos que $E(X) = 25$ y $\mathbf{V}ar(X) = 50$ de modo que

$$\begin{aligned} P(X \leq 34{,}384) &= P(\frac{X-25}{\sqrt{50}} \leq \frac{34{,}38225}{\sqrt{50}}) \\ &= P(Z \leq 1{,}33) \\ &= 0{,}9066. \end{aligned}$$

■

Definición 4.2. Sean $X_1, \ldots, X_{\mathbf{n}}$ variables aleatorias normales independientes con $\mathbf{E}(X_j) = \mu_j$ y varianza $\mathbf{V}ar(X_j) = \sigma^2, j = 1, 2, \ldots, \mathbf{n}$. También, sea $Y = \sum_{j=1}^{\mathbf{n}} X_i/\sigma_j^2$. La variable aleatoria Y se dice que se distribuye en forma chi-cuadrado *no-central* con parámetro de *no-centralidad* $\sum_{j=1}^{\mathbf{n}} \mu_j/\sigma^2$ y $\mathbf{n}$ grados de libertad. Se escribe como $Y \sim \chi^2(\mathbf{n}, \delta)$, donde $\delta = \sum_{j=1}^{\mathbf{n}} \mu_j/\sigma^2$.

La función de densidad de probabilidad de una variable aleatoria $\chi^2(\mathbf{n}, \delta)$ se puede demostrar que ses igual a

$$f_{\mathbf{n}}(y, \delta) = \begin{cases} \pi^{-\frac{1}{2}} 2^{-\mathbf{n}/2} \exp\left\{-\frac{1}{2}(\delta + y)\right\} y^{(\mathbf{n}-2)/2} & \\ \qquad \sum_{j=0}^{\infty} \frac{(\delta y)^j (\Gamma(j+1/2))}{(2j)!\Gamma(j1/2n)} & ; \quad y > 0 \\ \qquad 0 & ; \quad y \leq 0 \end{cases} \tag{4.3}$$

donde $\delta = \sum_{j=1}^{\mathbf{n}} \mu_j/\sigma^2$. Haciendo $\delta = 0$, observamos que ecuación (4.3) se reduce a la función de densidad de probabilidad $\chi^2(\mathbf{n})$ central.

Definición 4.3. Considere que la variable aleatoria X se distribuye en forma normal estándar ($X \sim \mathcal{N}(0, 1)$) y la variable aleatoria Y se distribuye en forma de distribución chi-cuadrado con $\mathbf{n}$ grados de libertad ($Y \sim \chi^2(\mathbf{n})$) y además, las variables X y Y son independientes. Entonces el estadístico

$$T = \frac{X}{\sqrt{Y/\mathbf{n}}} \tag{4.4}$$

se dice que tiene una distribución-t con $\mathbf{n}$ grados de libertad y se escribe $T \sim t(\mathbf{n})$.

Teorema 4.3. La función de densidad de probabilidad del estadístico definido en la ecuación (4.4) está dado por

$$f_{\mathbf{n}}(t) = \frac{\Gamma(\mathbf{n}+1)/2}{\Gamma(\mathbf{n}/2)\sqrt{\mathbf{n}\pi}}\left(1+\frac{t^2}{\mathbf{n}}\right)^{-\frac{1}{2}(\mathbf{n}+1)} \qquad -\infty < t < \infty \tag{4.5}$$

Observación 4.2. Para $\mathbf{n}=1, T$ es una variable aleatoria Cauchy. Por consiguiente se asumirá $\mathbf{n} > 1$. Para cada $\mathbf{n}$, se tendrá una función de densidad de probabilidad diferente. Análogo a la distribución normal,la distribución-t es importante en la en la teoría del estudio de los estadísticos y en consecuencia está tabulada.

Observación 4.3. La función de densidad de probabilidad $f_{\mathbf{n}}$ es simétrico en t, y cuando $f_{\mathbf{n}}(t) \longrightarrow 0$ cuando $t \longrightarrow +\infty$. Para $\mathbf{n}$ grande, la distribución-t es bastante cercano a la distribución normal. En efecto, $(1+t^2/\mathbf{n})^{-(\mathbf{n}+1)/2} \longrightarrow e^{-t^2/2}$ cuando $\mathbf{n} \longrightarrow \infty$. Sin embargo, para $\mathbf{n}$ pequeño, la variable T se desvía considerablemente de la normal. De hecho, $P(|T| \geq t_0) \geq P(|Z| \geq t_0), \quad Z \sim \mathcal{N}(0,1)$, esto es, existe mayor probabilidad en la cola de la distribución-t que en la cola de la distribución normal estándar. En lo que sigue se escribirá $t(\mathbf{n}, \alpha/2)$ para el valor de T para la cual

$$P(|T| > t(\mathbf{n}, \alpha/2)) = \alpha. \tag{4.6}$$

Teorema 4.4. Considere que $X \sim t(\mathbf{n}), \mathbf{n} > 1$. Entonces $E(X^r)$ existe para $r < \mathbf{n}$ y está dado por

$$E(X^r) = \begin{cases} 0 & \text{si} \quad r < \mathbf{n} \text{ es impar} \\ n^{\frac{r}{2}} \frac{\Gamma[\frac{(r+1)}{2}]\Gamma[\frac{(\mathbf{n}-r)}{2}]}{\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})}. & \text{si} \quad r < \mathbf{n} \text{ es par} \end{cases} \tag{4.7}$$

Corolario 4.2. Si $\mathbf{n} > 2$, $\mathbf{E}(X) = 0$ y $\mathbf{V}ar(X) = \mathbf{n}/(\mathbf{n}-2)$.

Observación 4.4. Si en la definición (4.3) tomamos que $X \sim \mathscr{N}(\mu, \sigma^2), Y/\sigma^2 \sim \chi^2(n)$, además X y Y son independientes

$$T = \frac{X}{\sqrt{\frac{Y}{\mathbf{n}}}}$$

se dice que tiene una distribución-t no-central con parámetro (llamado también parámetro de no-centralidad) $\delta = \mu/\sigma$ y n grados de libertad. La función de densidad de probabilidad de una distribución-t no-central está dado por

$$f_{\mathbf{n}}(t, \delta) = \frac{\mathbf{n}^{\mathbf{n}/2} e^{-\delta^2/2}}{\sqrt{\pi}\Gamma(\frac{\mathbf{n}}{2})(\mathbf{n}+t^2)^{(\mathbf{n}+1)/2}} \sum_{s=0}^{\infty} \Gamma\left(\frac{\mathbf{n}+s+1}{2}\right)\left(\frac{\delta^s}{s!}\right)\left(\frac{2t^2}{\mathbf{n}+t^2}\right)^{s/2} \tag{4.8}$$

si hacemos $\delta = 0$ en la ecuación (4.8), obtenemos la función de densidad de probabilidad $f_{\mathbf{n}}$ de una distribución-t central dada en la ecuación (4.5).

Si T tiene una distribución-t no-central, con n grados de libertad y parámetro de no-centralidad δ, entonces, la esperanza y la varianza están dadas por

$$\mathbf{E}(T) = \delta \frac{\Gamma\left[\frac{(\mathbf{n}-1)}{2}\right]}{\Gamma\left[\frac{\mathbf{n}}{2}\right]} \sqrt{\frac{\mathbf{n}}{2}}, \quad \mathbf{n} > 1 \tag{4.9}$$

y

$$\mathbf{V}ar(T) = \frac{\mathbf{n}(1+\delta^2)}{\mathbf{n}-2} - \frac{\delta^2 \mathbf{n}}{2}\left(\frac{\Gamma\left[\frac{(\mathbf{n}-1)}{2}\right]}{\Gamma\left[\frac{\mathbf{n}}{2}\right]}\right)^2, \quad \mathbf{n} > 2 \tag{4.10}$$

Definición 4.4. Sean X y Y variables aleatorias independientes chi-cuadras con **m** y **n** grados de libertad respectivamente. Entonces se dice que la variable aleatoria

$$F = \frac{X/\mathbf{m}}{Y/\mathbf{n}} \tag{4.11}$$

tiene distribución-F con $(\mathbf{m}, \mathbf{n})$ grados de libertad, y se escribe $F \sim F(\mathbf{m}, \mathbf{n})$.

Teorema 4.5. La función de densidad de probabilidad del estadístico-F dada en la ecuación (4.11) está dada por,

$$g(f) = \begin{cases} \frac{\Gamma\left[\frac{(\mathbf{m}+\mathbf{n})}{2}\right]}{\Gamma\left(\frac{\mathbf{m}}{2}\right)\Gamma\left(\frac{\mathbf{n}}{2}\right)} \left(\frac{\mathbf{m}}{\mathbf{n}}\right) \left(\frac{\mathbf{m}}{\mathbf{n}} f\right)^{\frac{\mathbf{m}}{2}-1} \left(1 + \frac{\mathbf{m}}{\mathbf{n}} f\right)^{-\frac{(\mathbf{m}+\mathbf{n})}{2}} & f > 0, \\ 0, & f \leq 0 \end{cases} \tag{4.12}$$

Observación 4.5. Si $X \sim F(\mathbf{m}, \mathbf{n})$, entonces $1/X \sim F(\mathbf{n}, \mathbf{m})$. Si hacemos $m = 1$, entonces $F = [t(\mathbf{n})]^2$ de modo que $F(1, \mathbf{n})$ y $t^2(\mathbf{n})$ tienen la misma distribución. También se continúa, que si $Z \sim \mathscr{C}$ [la cual es la misma como $t(1)$], $Z^2 \sim F(1, 1)$.

Observación 4.6. Como es natural, escribimos $F(\mathbf{m}, \mathbf{n}; \alpha)$ para el porcentaje de puntos supriores de la distribución $F(\mathbf{m}, \mathbf{n})$, esto es,

$$P\left(F(\mathbf{m}, \mathbf{n}) > F(\mathbf{m}, \mathbf{n}; \alpha)\right) == \alpha \tag{4.13}$$

De la observación 4.5, se tiene la siguiente relación

$$F(\mathbf{m}, n; 1-\alpha) = \frac{1}{F(\mathbf{n}, \mathbf{m}; \alpha)} \tag{4.14}$$

Por tanto es suficiente tabular valores de F que sean mayores o iguales a 1 $(F \geq)$.

Teorema 4.6. Si $X \sim F(\mathbf{m}, \mathbf{n})$. Entonces, para $k > 0$, entero,

$$\mathbf{E}(X^k) = \left(\frac{\mathbf{n}}{\mathbf{m}}\right)^k \frac{\Gamma\left[k + \frac{\mathbf{m}}{2}\right]\Gamma\left[\frac{\mathbf{n}}{2} - k\right]}{\Gamma\left[\frac{\mathbf{m}}{2}\right]\Gamma\left[\frac{\mathbf{n}}{2}\right]} \qquad \mathbf{n} > 2k \tag{4.15}$$

En particular,

$$\mathbf{E}(X) = \frac{\mathbf{n}}{\mathbf{m}}, \quad \mathbf{n} > 2, \tag{4.16}$$

y

$$\mathbf{V}ar(X) = \frac{\mathbf{n}^2(2\mathbf{m} + 2\mathbf{n} - 4)}{\mathbf{m}(\mathbf{n} - 2)^2(\mathbf{n} - 4)}, \quad \mathbf{n} > 4. \tag{4.17}$$

Teorema 4.7. Si $X \sim F(\mathbf{m}, \mathbf{n})$, entonces, entonces $Y = 1/[1 + (\mathbf{m}/\mathbf{n})X]$ es $B(\mathbf{n}/2, \mathbf{m}/2)$. Consecuentemente, para cada $x > 0$

$$F_X(x) = 1 - F_Y\left[\frac{1}{1 + (\mathbf{m}/\mathbf{n})x}\right]$$

Demostración. Se deja como ejercicio □

Si en la definición (4.4) hacemos que $X \sim \chi^2(\mathbf{n}, \delta)$ se obtiene la variable aleatoria F no-central.

Definición 4.5. Sea $X \sim \chi^2(\mathbf{n}, \delta)$ y $Y \sim \chi^2(\mathbf{n})$, y sean X yY independientes. Entonces la variable aleatoria

$$F = \frac{X/\mathbf{m}}{Y/\mathbf{n}} \tag{4.18}$$

se dice que tiene un a distribución-F no-central con grados de libertad $(\mathbf{m}, \mathbf{n})$ y parámetro de no-centralidad δ.

Se puede demostrar que la función de densidad de probabilidad de la variable aleatoria F definida en la relación (4.18) está dada por

$$g(f,m,n,\delta)=\begin{cases}\frac{\mathbf{m}^{\mathbf{m}/2}\mathbf{n}^{\mathbf{n}/2}}{\Gamma(\frac{\mathbf{n}}{2})}e^{-\delta/2}f^{(\mathbf{m}/2)-1} \\ \times\displaystyle\sum_{j=0}^{\infty}\frac{(\delta\mathbf{m}f)^{j}\Gamma[\frac{(\mathbf{m}+\mathbf{n})}{2}+j]}{j!\Gamma(\frac{\mathbf{m}}{2}+j)(\mathbf{m}f+\mathbf{n})^{(\mathbf{m}+\mathbf{n})/2+j}} & \text{si } f>0 \\ 0, & \text{si } f\leq 0.\end{cases} \tag{4.19}$$

Substituyendo $\delta=0$, obtenemos la función de densidad de $F(m,n)$ central dada en la ecuación (4.12).

5. Matriz de Helmert

En álgebra lineal y teoría de matrices hay muchas matrices especiales e importantes. Por ejemplo, la conocida matriz de Helmert es una de ellas. Una matriz de Helmert de orden $\mathbf{n}$ es una matriz cuadrada que fue introducida por H. O. Lancaster en 1965. Por lo general, la matriz de Helmert se usa en estadísticas matemáticas para el análisis de varianza (ANOVA). También se puede utilizar en procesos estocásticos. De hecho, sabe que en la teoría moderna de la probabilidad y los sistemas dinámicos, las matrices estocásticas son matrices reales no negativas que se utilizan para mostrar las probabilidades de transición.

$$A = \begin{bmatrix} \frac{1}{\sqrt{\mathbf{n}}} & \frac{1}{\sqrt{\mathbf{n}}} & \frac{1}{\sqrt{\mathbf{n}}} & \cdots & \frac{1}{\sqrt{\mathbf{n}}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{2\times 3}} & \frac{1}{\sqrt{2\times 3}} & -\frac{2}{\sqrt{2\times 3}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \frac{1}{\sqrt{(\mathbf{n}-1)\mathbf{n}}} & \frac{1}{\sqrt{(\mathbf{n}-1)\mathbf{n}}} & \frac{1}{\sqrt{(\mathbf{n}-1)\mathbf{n}}} & \cdots & -\frac{(\mathbf{n}-1)}{\sqrt{(\mathbf{n}-1)\mathbf{n}}} \end{bmatrix}$$

$$AX = \begin{bmatrix} \frac{1}{\sqrt{\mathbf{n}}} & \frac{1}{\sqrt{\mathbf{n}}} & \frac{1}{\sqrt{\mathbf{n}}} & \cdots & \frac{1}{\sqrt{\mathbf{n}}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{2\times 3}} & \frac{1}{\sqrt{2\times 3}} & -\frac{2}{\sqrt{2\times 3}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \frac{1}{\sqrt{(\mathbf{n}-1)\mathbf{n}}} & \frac{1}{\sqrt{(\mathbf{n}-1)\mathbf{n}}} & \frac{1}{\sqrt{(\mathbf{n}-1)\mathbf{n}}} & \cdots & -\frac{(\mathbf{n}-1)}{\sqrt{(\mathbf{n}-1)\mathbf{n}}} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_{\mathbf{n}} \end{bmatrix}$$

lo que resulta

$$AX = \begin{bmatrix} \sqrt{\mathbf{n}}\overline{X} \\ \frac{X_1 - X_2}{2} \\ \frac{X_1 + X_2 - 2X_3}{\sqrt{2\times 3}} \vdots \\ \frac{X_1 + X_2 + \cdots + X_{\mathbf{n}-1} - (\mathbf{n}-1)X_{\mathbf{n}}}{\sqrt{(\mathbf{n}-1)\mathbf{n}}} \end{bmatrix}$$

denotando los componentes de ASX por

$$\begin{aligned}
U_0 &= \sqrt{\mathbf{n}}\overline{X} \\
U_1 &= \frac{X_1 - X_2}{2} \\
&\vdots \\
U_{\mathbf{n}-1} &= \frac{X_1 + X_2 + \cdots + X_{\mathbf{n}-1} - (\mathbf{n}-1)X_{\mathbf{n}}}{\sqrt{(\mathbf{n}-1)\mathbf{n}}}
\end{aligned}$$

El vector aleatorio U se puede escribir como

$$U = \begin{bmatrix} U_0 \\ U_1 \\ \vdots \\ U_{\mathbf{n}-1} \end{bmatrix} = AX$$

El vector aleatorio $U \sim \mathscr{N}(\boldsymbol{I}_{\mathbf{n}}\mu, \sigma^2 \boldsymbol{I}_{\mathbf{n}})$, donde $\boldsymbol{I}_{\mathbf{n}}$ es la matriz identidad de orden $\mathbf{n}$ y $\mu = \mathbf{E}(X_j), j = 1, 2, \ldots, \mathbf{n}$.
Definiendo

$$U_j = \sum_{k=1}^{j} \frac{X_k - jX_{j+1}}{\sqrt{j(j+1)}}, \; j = 1, 2, \ldots, (\mathbf{n}-1).$$

Ejercicio 1

Teniendo en cuenta A, el vector X, y U definida antes

(a) Escriba $U = AX$ para $\mathbf{n} = 4$.

(b) Demuestre directamente que para $j \geq 1$, $U_j \sim \mathscr{N}(0, \sigma^2)$. *sugerencia* Encuentre la función característica de U_j.

Ejercicio 2

En este ejercicio solo use las propiedades algebraicas de S^2,A, $\mathbf{X}$, U; no use la distribución normal.

(a) Demuestre $A'A = I = AA'$

(b) Demuestre que A(1,1,..., 1)$'\boldsymbol{I}_{\mathbf{n}} = (\sqrt{\mathbf{n}}, 0, \ldots, 0)'$.

(c) Demuestre que

$$\begin{aligned}\sum_{j=1}^{n} X_j^2 &= X'X = XA'AX \\ &= U'U = \mathbf{n}(\overline{X})^2 + \sum_{j=1}^{\mathbf{n}-1} U_J^2.\end{aligned}$$

sugerencia: intente para $n = 4$ primero

d Demuestre que

$$\begin{aligned}(\mathbf{n}-1)S^2 &= \sum_{j=1}^{\mathbf{n}} (X_j - \overline{X})^2 \\ &= \sum_{j=1}^{\mathbf{n}} X_j^2 - \mathbf{n}\overline{X}^2 = \sum_{j=1}^{\mathbf{n}-1} U_J^2.\end{aligned}$$

La derivación de la distribución de $(n-1)S^2$ se completará si demuestra que $U_1, U_2 \ldots, U_{\mathbf{n}-1}$ sean independientes y luego tome $Z_j = \frac{U_j}{\sigma}$

Ejercicio 3

Demuestre que los $U_j, j = 0, 1, 2, \ldots, (\mathbf{n} - 1)$, son independientes. Luego

$$\overline{X} = \frac{U_0}{\sqrt{\mathbf{n}}} \quad \text{y} \quad S^2 = \frac{\sum_{j}^{\mathbf{n}} (X_j - \overline{X})^2}{\mathbf{n} - 1}$$

son funciones de variables aleatorias independientes y por lo tanto son independientes uno del otro.

Ejercicio 4

Encuentre la función característica de de U

Para todo $a, b \in \mathbb{R} : a \leq b$ y para toda variable aleatoria X continua se tiene:

$$\begin{aligned} P(a \leq X \leq b) &= P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) \\ &= F_X(b) - F_X(a), \quad \text{donde } F_X \text{ es la fd} \end{aligned}$$

Si $X \sim \mathcal{N}(\mu, \sigma^2)$ entonces se sabe que $\dfrac{\sqrt{\mathbf{n}}(\overline{X} - \mu)}{\sigma} \sim \mathcal{N}(0, 1)$. Por lo que:

$$\begin{aligned} P(a \leq \overline{X} \leq b) &= P(a < \overline{X} \leq b) = P(a \leq \overline{X} < b) = P(a < \overline{X} < b) \\ &= \Phi\left(\frac{\sqrt{n}(b - \mu)}{\sigma}\right) - \Phi\left(\frac{\sqrt{n}(a - \mu)}{\sigma}\right) \end{aligned}$$

donde Φ es la distribución normal estándar.

Ejemplo 5.1. Sea $X_1, X_2, \ldots, X_{16}$ una muestra aleatoria de una distribución normal, $\mathcal{N}(77, 25)$: Calcular:

(a) $P(77 < \overline{X} < 79{,}5)$

(b) $P(74{,}2 < \overline{X} < 78{,}4)$

■

Solución del item (a)

$$\begin{aligned} P(77 < \overline{X} < 79{,}5) &= \Phi\left(\frac{\sqrt{16}(79{,}5-77)}{5}\right) - \Phi\left(\frac{\sqrt{16}(77-77)}{5}\right) \\ &= \Phi(2) - \Phi(0) \\ &= 0{,}97725 - 0{,}50000 = 0{,}47725 \end{aligned}$$

Solución del item (b)

$$\begin{aligned} P(74{,}2 < \overline{X} < 78{,}4) &= \Phi\left(\frac{\sqrt{16}(78{,}4-77)}{5}\right) - \Phi\left(\frac{\sqrt{16}(74{,}2-77)}{5}\right) \\ &= \Phi(1{,}12) - \Phi(-2{,}24) \\ &= 0{,}86864 - 0{,}01255 = 0{,}85609 \end{aligned}$$

Ejercicio 5

El número de libros encuadernados diariamente por una máquina automática es una variable aleatoria cuya distribución se desconoce, con una desviación típica de 16 libros por día. Si se selecciona una muestra aleatoria de 49 días, hallar la probabilidad de que el número medio de libros encuadernados durante esos días (la media muestral) se encuentre a lo más 3 libros de la verdadera media poblacional.

Solución 5

Como no se especifica la distribución del número de libros y $n = 49 > 30$, podemos utilizar el TCL. para hallar la probabilidad solicitada. Tenemos

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

La probabilidad que pide es el siguiente

$$\begin{aligned}
P(\left|\overline{X}-\mu\right| \leq 3) &= P(-3 \leq \overline{X}-\mu \leq 3) \\
&= P\left(\frac{-3}{16/\sqrt{49}} \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{3}{16/\sqrt{49}}\right) \\
&= \Phi(1{,}3125) - \Phi(-1{,}3125) \\
&= 0{,}81064
\end{aligned}$$

Solución 6

Con referencia al ejercicio 5. Determinar el tamaño de la muestra para que la media muestral se encuentre a lo más a 3 libros de la media poblacional con una probabilidad de 0,95.

Solución 6

Ahora se tiene que verificar que

$$P(\left|\overline{X}-\mu\right| \leq 3) = P(-3 \leq \overline{X}-\mu \leq 3) = 0{,}95$$

Estandarizando se tiene

$$\begin{aligned}
P\left(\frac{-3}{16/\sqrt{49}} \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{3}{16/\sqrt{49}}\right) &= P(-0{,}1875\sqrt{n} \leq Z \leq 0{,}1857\sqrt{n}) \\
&= \Phi(0{,}1875\sqrt{n}) - \Phi(-0{,}1875\sqrt{n}) = 0{,}95 \\
2\Phi(0{,}8175\sqrt{n}) - 1 &= 0{,}95 \\
\Phi(0{,}1875\sqrt{n}) &= 0{,}975 \\
0{,}1875\sqrt{n} &= \Phi^{-1}(0{,}975) \\
0{,}1875\sqrt{n} &= 1{,}95996 \\
n &\approx 110
\end{aligned}$$

Ejercicios propuestos

1. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de $\mathcal{N}(\mu, \sigma^2)$ y $\overline{X}$ y S^2, respectivamente, la media muestral y la varianza muestral. Sea $X_{n+1} \sim \mathcal{N}(\mu, \sigma^2)$, y suponga que $X_1, X_2, \ldots, X_n$, X_{n-1} son independientes. Encuentre la distribución de muestreo de $\left[\frac{(X_{n+1}-\overline{X})}{S/\sqrt{\mathbf{n}/(n+1)}}\right]$

2. Sean $X_1, X_2, \ldots, X_{\mathbf{m}}$ y $Y_1, Y_2, \ldots, Y_n$ muestras aleatorias independientes de $\mathcal{N}(\mu_X, \sigma_X^2)$ y $\mathcal{N}(\mu_Y, \sigma_Y^2)$, respectivamente. Además, sean α y β dos números reales fijos. Si $\overline{X}$, $\overline{Y}$ denotan las medias muestrales correspondientes, ¿cuál es la distribución muestral de
$$\frac{\alpha(\overline{X}-\mu_X)+\beta(\overline{Y}-\mu_Y)}{\sqrt{\frac{(m-1)S_X^2+(n-1)S_Y^2}{m+n-2}}\sqrt{\frac{\alpha^2}{m}+\frac{\beta^2}{n}}}$$
donde S_X^2 y S_Y^2respectivamente denotan las varianzas muestrales de ls $X's$ y de las Y'.

3. Sea $X_1, X_2, \ldots, X_{\mathbf{n}}$ una muestra aleatoria de $\mathcal{N}(\mu, \sigma^2)$ y sea k es un número entero entero positivo. Encuentra $\mathbf{E}(S^{2k})$. En particular, encuentre $\mathbf{E}(S^2)$ y $\mathbf{V}ar(S^2)$.

4. Se toma una muestra aleatoria de 5 de una población normal con una media de 2,5 y una varianza $\sigma^2 = 36$.

   (a) Encuentre la probabilidad de que la varianza muestral se encuentre entre 30 y 44.

   (b) Encuentre la probabilidad de que la media muestral se encuentre entre 1.3 y 3.5, mientras que la varianza muestral se encuentra entre 30 y 44.

5. Se observó que la vida media de una muestra de 10 bombillas era de 1327 horas con una desviación estándar de 425 horas. Una segunda muestra de 6 bombillas elegidas de un lote diferente mostró una vida media de 1215 horas con una desviación estándar de 375 horas. Si se supone que las medias de los dos lotes son las mismas, ¿qué tan probable es la observación? diferencia entre las dos medias de muestra?

6. Sean S_X^2 y S_Y^2 las varianzas muestrales de dos muestras independientes de tamaños $n_x = 5$ y $n_Y = 4$ de dos poblaciones que tienen la misma varianza desconocida σ^2. Encuentre (aproximadamente) la probabilidad de que $\frac{S_X^2}{S_Y^2} < 1/5{,}2$ o $> 6{,}25$.

7. Sean $X_1, X_2, \ldots, X_\mathbf{n}$ variables aleatorias iid que tienen un $\mathbf{E}|X_1|^4$ finito y sean $\overline{X}$ y S^2 la media y la varianza de la muestra. Exprese $\mathbf{E}(\overline{X}^3)$, $\mathbf{C}ov(\overline{X}, S^2)$ y $\mathbf{V}ar(S^2)$ en términos de $\mu_k = \mathbf{E}(X_1^k)$, $k = 1, 2, 3, 4$. Encuentre una condición bajo la cual $\overline{X}$ y S^2 sean no correlacionados

8. Sean $X_1, X_2, \ldots, X_\mathbf{n}$ variables aleatorias iid que tienen la distribución gamma $\Gamma(\alpha, \beta_X)$ y $Y_1, Y_2, \ldots, Y_\mathbf{n}$. variables aleatorias iid que tienen la distribución gamma $\Gamma(\alpha, \beta_Y)$, donde $\alpha > 0$, $\beta_X > 0$ y $\beta_Y > 0$. Suponga que las X_i y las Y_i son independientes. Derive la distribución del estadístico $\dfrac{\overline{X}}{\overline{Y}}$, donde $\overline{X}$ y $\overline{Y}$ son las medias muestrales basadas en las X_i y en las Y_i respectivamente.

9. Sean $X_1, X_2, \ldots, X_\mathbf{n}$ variables aleatorias iid que tienen la distribución exponencial $\exp(\mu, \theta)$, $\mu \in \mathbb{R}$ y $\theta > 0$. Demuestre que el estadístico de orden más mínimo, $X_{(1)}$, tiene la dis-

tribución exponencial $\exp(\mu, \theta/\mathbf{n})$ y que $2\sum_{I=1}^{\mathbf{n}}(X_i - X_{(1)})/\theta$ tiene la distribución chi-cuadrado $\chi^2(2\mathbf{n}-2)$.

6. Distribuciones de $\overline{X}$ para algunas Poblaciones no Normales

Hasta ahora hemos considerado estadísticas que eran funciones de variables aleatorias con las mismas distribuciones normales. Ahora investigamos algunas de la distribución de medias muestrales de muestras que provienen de poblaciones que no son normales.

6.1. Distribución de $\overline{X}$ de la distribución Binomial

Sea $X_1, X_2, \ldots, X_m$ una muestra aleatoria de una población con distribución Binomial de parámetros $\mathbf{n}$ y θ. La función de probabilidad de la variable X_j es

$$\mathbf{P}(X_j = r) = \binom{\mathbf{n}}{r}\theta^r(1-\theta)^{\mathbf{n}-r}, \quad \begin{array}{l} r = 0, 1, 2, \ldots, \mathbf{n} \\ j = 1, 2, \ldots, \mathbf{m}. \end{array}$$

donde $0 < \theta < 1$. Se desea determinar la distribución de la media muestral de estas variables aleatorias.

$$\overline{X} = \frac{1}{\mathbf{m}}\sum_{j=1}^{\mathbf{m}} X_j$$

Para ello utilizamos la función característica; se sabe que la función característica de la $\overline{X}$, de una muestra de tamaño $\mathbf{n}$ está dada por

$$\Psi_{\overline{X}}(t) = \left[\Psi_X\left(\frac{t}{\mathbf{n}}\right)\right]^{\mathbf{n}}$$

donde Ψ_X es la función característica de la población. En caso, la binomial. La función característica de esta población es

$$\Psi_X(t) = (1 - \theta + \theta e^{it})^{\mathbf{n}}$$

Luego, la función característica de $\overline{X}$ es

$$\Psi_{\overline{X}}(t) = (1 - \theta + \theta e^{i\frac{t}{\mathbf{m}}})^{\mathbf{nm}} \tag{6.1}$$

La expresión (6.1) es la función característica de una variable aleatoria con función de distribución binomial modificado: $\overline{X}$ puede tomar los valores

$$\overline{X} = 0, \frac{1}{\mathbf{m}}, \frac{2}{\mathbf{m}}, \ldots, \frac{\mathbf{nm}}{\mathbf{m}} = \mathbf{n},$$

y la función de probabilidad de la media muestra , $\overline{X}$, es

$$\mathbf{P}(\overline{X} = \frac{k}{\mathbf{m}}) = \binom{\mathbf{nm}}{k}\theta^k(1-\theta)^{\mathbf{nm}-k}, k = 0, 1, 2, \ldots, \mathbf{nm}$$

de donde obtiene que

$$\mathbf{E}(\overline{X}) = \mathbf{E}\left(\frac{k}{\mathbf{m}}\right) = \frac{1}{\mathbf{m}}\mathbf{E}(k) = \frac{1}{\mathbf{m}}\mathbf{nm}\theta = \mathbf{n}\theta$$

y

$$\mathbf{V}ar(\overline{X}) = \mathbf{V}ar\left(\frac{k}{\mathbf{m}}\right) = \frac{1}{\mathbf{m}^2}\mathbf{V}ar(k) = \frac{1}{\mathbf{m}^2}\mathbf{nm}\theta(1-\theta) = \frac{\mathbf{n}\theta(1-\theta)}{\mathbf{m}}$$

6.2. Distribución de $\overline{X}$ de la distribución Poisson

Ahora considere variables aleatorias independientes con la misma $X_j (j = 1, 2, \ldots, \mathbf{m})$ con la misma distribución de Poisson dada por

$$\mathbf{P}(X_j = r) = e^{-\lambda}\frac{\lambda^r}{r!}, \quad \begin{array}{l} r = 0, 1, 2, \ldots, \\ j = 1, 2, \ldots, \mathbf{m}. \end{array}$$

donde $\lambda > 0$.

Para hallar la distribución del estadístico $\overline{X}$ definido en la relación (6.1), observe que la función característica de la población Poisson es,

$$\Psi_X(t) = e^{\lambda(e^{it}-1)}.$$

Entonces la función característica de $\overline{X}$ es

$$\Psi_{\overline{X}}(t) = e^{\mathbf{m}\lambda(e^{i\frac{t}{m}}-1)}$$

La expresión anterior, es la función característica de una variable aleatoria con una modificación de la distribución Poisson; $\overline{X}$ puede tomar los valores

$$\overline{X} = 0, \frac{1}{\mathbf{m}}, \frac{2}{\mathbf{m}}, \frac{3}{\mathbf{m}}, \cdots$$

y la función de probabilidad de la media muestra , $\overline{X}$, es

$$\mathbf{P}\left(\overline{X} = \frac{k}{\mathbf{m}}\right) = e^{-\mathbf{m}\lambda}\frac{(\mathbf{m}\lambda)^k}{k!}, \; k = 0, 1, 2, \ldots$$

La media y varianza se puede hallar utilizando directamente la función característica. Así

$$\Psi'_{\overline{X}}(t) = \lambda i e^{\mathbf{m}\lambda(e^{i\frac{t}{m}}-1)}\left[e^{i\frac{t}{\mathbf{m}}}\right]$$

y

$$\Psi'_{\overline{X}}(0) = \lambda i = i\mathbf{E}(\overline{X}) \Rightarrow \mathbf{E}(\overline{X}) = \lambda$$

Ejercicio 7

Encuentre la distribución de la media muestral $\overline{X}$ de una muestra aleatoria simple de tamaño $\mathbf{n}$ extraída de una población en la que característica X tiene la distribución Gamma con parámetros α y β.

Ejercicio 8

Sea $\overline{X}$ la media muestral de una muestra aleatoria simple de tamaño $\mathbf{n}$ extraída de una población en la cual la característica X tiene la distribución uniforme dada por

$$f(x) = \begin{cases} 1 & \text{si} \quad x \in [0,1] \\ 0 & \text{si} \quad x \notin [0,1] \end{cases}$$

Pruebe la función de densidad $g(x)$ de $\overline{X}$ es para $j = 0, 1, 2, \ldots, n-1$ es de la forma

$$g_{\overline{X}}(x) = \frac{\mathbf{n}^{\mathbf{n}}}{(\mathbf{n}-1)!} \sum_{k=0}^{j} (-1)^k \binom{\mathbf{n}}{k} \left(x - \frac{k}{\mathbf{n}}\right)^{\mathbf{n}-1}, \left(\frac{j}{\mathbf{n}} \leq x \leq \frac{j+1}{\mathbf{n}}\right)$$

7. Muestreo en Poblaciones Normales: Distribución de la diferencia de medias muestrales con varianza conocidas

En muchas situaciones surge la necesidad de comparar las medias muestrales de dos poblaciones distintas. Por ejemplo supongamos que estamos interesados en comparar los tiempos medios de duración de dos tipos de tubos fluorescente. La fabricación de ambos tipos de tubos de fluorescentes se realiza por empresas distintas y con diferentes procesos de fabricación. Por lo tanto, los tubos producidos por cada empresa tendrán una distribución diferente, una de la otra, de los tiempos de duración de los tubos.

Teorema 7.1. Sean $X_1, X_2, \ldots, X_{\mathbf{m}}$ y $Y_1, Y_2, \ldots, Y_{\mathbf{n}}$ dos muestras aleatorias, cada una de distribuciones normales independientes con medias desconocidas μ_x y μ_y, pero varianzas conocidas σ_x^2 y σ_y^2, respectivamente. Entonces

$$\overline{X} - \overline{Y} \sim \mathscr{N}\left(\mu_x - \mu_y, \frac{\sigma_x^2}{\mathbf{m}} + \frac{\sigma_y^2}{\mathbf{n}}\right)$$

Es decir, la diferencia de medias muestrales, $\overline{X} - \overline{Y}$, se distribuye en forma normal con $\mathbf{E}(\overline{X} - \overline{Y}) = \mu_x - \mu_y$ y varianza. $\mathbf{V}ar(\overline{X} - \overline{Y}) = \sigma_{\bar{x}-\bar{y}}^2 = \frac{\sigma_x^2}{\mathbf{m}} + \frac{\sigma_y^2}{\mathbf{n}}$.

Además,

$$Z = \frac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{\mathbf{m}} + \frac{\sigma_y^2}{\mathbf{n}}}} \sim \mathscr{N}(0, 1).$$

Prueba. Como las muestras son independientes, también lo son las medias muestrales $\overline{X}, \overline{Y}$. Luego por la propiedad del función característica resulta

$$\Psi_{\overline{X}-\overline{Y}}(t) = \Psi_{\overline{X}+(-\overline{Y})}(t) = \Psi_{\overline{X}}(t)\Psi_{-\overline{Y}}(t) = \Psi_{\overline{X}}(t)\Psi_{\overline{Y}}(-t) = \Psi_{\overline{X}}(t)\overline{\Psi_{\overline{Y}}(t)}$$

y como $\Psi_{\overline{X}}(t) = e^{it\mu_x + \frac{\sigma_x^2}{\mathbf{m}}\frac{t^2}{2}}$ y $\overline{\Psi_{\overline{Y}}(t)} = e^{-it\mu_y + \frac{\sigma_y^2}{\mathbf{n}}\frac{t^2}{2}}$. Entonces la función característica de la diferencia de medias muestrales resulta

$$\Psi_{\overline{X}-\overline{Y}}(t) = e^{it\mu_x + \frac{\sigma_x^2}{\mathbf{m}}\frac{t^2}{2}} e^{-it\mu_y + \frac{\sigma_y^2}{\mathbf{n}}\frac{t^2}{2}} = e^{it(\mu_x - \mu_y) + \left(\frac{\sigma_x^2}{\mathbf{m}} + \frac{\sigma_y^2}{\mathbf{n}}\right)\frac{t^2}{2}}.$$

Luego por el teorema de la unicidad de la función característica se concluye que

$$\overline{X} - \overline{Y} \sim \mathscr{N}\left(\mu_x - \mu_y, \frac{\sigma_x^2}{\mathbf{m}} + \frac{\sigma_y^2}{\mathbf{n}}\right).$$

$\square$

Ejemplo 7.1. Analizando los sueldos de los empleados de dos empresas se deduce que en la empresa A el salario medio mensual es de 2900 nuevos soles con una varianza de 250 (nuevos soles)^2, y en la empresa B el salario medio es de 2508 nuevos soles con una con una varianza de 300 (nuevos soles)^2. Si se toma una muestra aleatoria de 25 personas de la empresa A y de 36 personas de la empresa B. Determinar la probabilidad de que la muestra procedente de la empresa A tenga un salario medio que sea al menos 400 nuevos soles superior al salario medio de la empresa B.

Solución:

La información que se tiene es la siguiente:

Población A: $\mu_x = 2900$ $\sigma_x^2 = 250$ $m = 25$

Población B: $\mu_y = 2508$ $\sigma_y^2 = 300$ $m = 36$

Por Teorema 1, se sabe que

$$\overline{X} - \overline{Y} \sim \mathcal{N}\left(2900 - 2508, \frac{250}{25} + \frac{300}{36}\right)$$

$$\overline{X} - \overline{Y} \sim \mathcal{N}(392, 18{,}33)$$

Luego, tenemos

$$\begin{aligned}
P(\overline{X} - \overline{Y} > 400) &= P\left(\frac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{\mathbf{m}} + \frac{\sigma_y^2}{\mathbf{n}}}} > \frac{400 - 392}{\sqrt{18{,}33}}\right) \\
&= P\left(Z > \frac{8}{4{,}2814}\right) \\
&= P(z > 1{,}8685) = 1 - P(z > 1{,}8685) = 1 - \Phi(1{,}8685) \\
&= 0{,}03085
\end{aligned}$$

$P(\overline{X} - \overline{Y} > 400) = 0{,}03085$ ■

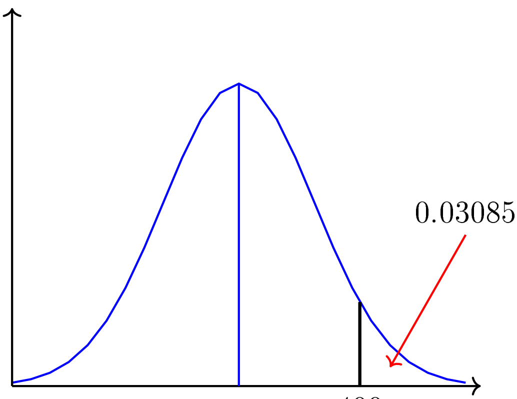

Figura 2.1: Representación gráfica de la muestral de la diferencia de medias muestrales.

7.1. Distribución de la diferencia de medias muestrales en Poblaciones normalmente distribuidas con varianzas desconocidas

Ahora se desea encontrar la distribución la diferencia de dos medias muestrales cuando las varianzas poblacionales son desconocidas. Sean $X_1, X_2, \ldots, X_{\mathbf{m}}$ y $Y_1, Y_2, \ldots, Y_{\mathbf{n}}$ dos muestras aleatorias, cada una de distribuciones normales independientes con medias desconocidas μ_x y μ_y y varianzas desconocidas σ_x^2 y σ_y^2, respectivamente. Se puede considerar dos casos, y a continuación tomar cada uno por turno:

(a) Ambas varianzas poblacionales son iguales, $\sigma_x^2 = \sigma_y^2 = \sigma^2$. En este caso, se asume que ambas muestras provienen de poblaciones que pueden tener varianzas iguales. Esto significa que se puede usar ambas muestras combinadas para estimar σ^2.

(b) Ambas varianzas poblacionales son desiguales, $\sigma_x^2 \neq \sigma_y^2$. En este caso, se asumimos que ambas muestras provienen de poblaciones que no tienen varianzas iguales, por lo que se debe estimar σ_x^2 y σ_y^2 por separado.

Caso (a) **:**$\sigma_x^2 = \sigma_y^2 = \sigma^2$**:**
Sea S_X^2 la varianza muestral de $X_1, X_2, \ldots, X_{\mathbf{m}}$, y sea S_Y^2 la varianza muestral de $Y_1, Y_2, \ldots, Y_{\mathbf{n}}$. Sabemos que S_X^2 es un estimador insesgado para σ_x^2 y que S_Y^2 es un estimador insesgado para σ_y^2. Como se asume que $\sigma_x^2 = \sigma_y^2 = \sigma^2$, entonces tanto S_X^2 como S_Y^2 son estimadores insesgados de σ^2.

La razón por la que este caso es interesante radica principalmente en la siguiente idea. Tenemos dos estimadores insesgados de σ^2. ¿Podríamos combinar estos estimadores de alguna manera para obtener otro estimador insesgado, que sea mejor que cualquiera de los dos estimadores insesgados originales? La respuesta a ésta pregunta es sí. Pero, ¿a qué nos referimos con mejor?

En esta situación, nos referimos a obtener otro estimador insesgado que tenga una varianza menor que cualquiera de los dos estimadores insesgados originales. Se sabe que

$$\frac{(\mathbf{m}-1)S_X^2}{\sigma_x^2} \sim \chi^2(\mathbf{m}-1) \quad \text{y} \quad \frac{(\mathbf{n}-1)S_Y^2}{\sigma_y^2} \sim \chi^2(\mathbf{n}-1)$$

y por la propiedad de la distribución χ^2 se tiene que,

$$\frac{(\mathbf{m}-1)S_X^2}{\sigma_x^2} + \frac{(\mathbf{n}-1)S_Y^2}{\sigma_y^2} \sim \chi^2(\mathbf{n}-1) \sim \chi^2(\mathbf{m}+\mathbf{n}-2)$$

y como las varianzas son iguales resulta

$$\frac{(\mathbf{m}-1)S_X^2}{\sigma^2} + \frac{(\mathbf{n}-1)S_Y^2}{\sigma^2} \sim \chi^2(\mathbf{n}-1) \sim \chi^2(\mathbf{m}+\mathbf{n}-2) \qquad (*)$$

Por otro lado, se sabe que

$$Z = \frac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{\mathbf{m}} + \frac{\sigma_y^2}{\mathbf{n}}}} \sim \mathscr{N}(0,1).$$

Nuevamente cono las varianzas son iguales resulta,

$$Z = \frac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_y)}{\sigma\sqrt{\frac{1}{\mathbf{m}} + \frac{1}{\mathbf{n}}}} \sim \mathscr{N}(0,1). \qquad (**)$$

y teniendo en cuenta, que las medias muestrales son independientes de las varianzas muestrales, por definición del estadístico "t" student, efectuando la división del de relación de $(**)$ entre la relación de $(*)$ se obtiene

$$\begin{aligned} T &= \frac{\dfrac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_y)}{\sigma\sqrt{\dfrac{1}{\mathbf{m}} + \dfrac{1}{\mathbf{n}}}}}{\sqrt{\dfrac{\dfrac{(\mathbf{m}-1)S_X^2}{\sigma^2} + \dfrac{(\mathbf{n}-1)S_Y^2}{\sigma^2}}{\mathbf{m}+\mathbf{n}-2}}} \sim t(\mathbf{m}+\mathbf{n}-2) \\ &= \frac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_y)}{S\sqrt{\dfrac{1}{\mathbf{m}} + \frac{1}{\mathbf{n}}}} \sim t(\mathbf{m}+\mathbf{n}-2) \end{aligned}$$

donde

$$S^2 = \frac{(\mathbf{m}-1)S_X^2 + (\mathbf{n}-1)S_Y^2}{\mathbf{m}+\mathbf{n}-2}$$

se llama **varianza ponderada** y es un estimador insesgado de σ^2.

Ejemplo 7.2. Se midió la proporción de consumo de oxígeno de dos grupos de hombres. Un grupo (X) entrenó regularmente durante un período de tiempo, y el otro (Y) entrenó de forma intermitente.

Las estadísticas calculadas a partir de los datos registrados se dan a continuación: Suponiendo

Grupo X :	$\mathbf{m} = 9$	$\mu_x = 43{,}71$	$s_X^2 = 34{,}5744$
Grupo Y :	$\mathbf{n} = 7$	$\mu_Y = 39{,}63$	$s_Y^2 = 58{,}9824$

distribuciones normales independientes para las lecturas, y que las varianzas de las poblaciones son iguales, encuentre la $P(\overline{X} > \overline{Y})$.

Solución:

$$\begin{aligned}
P(\overline{X} > \overline{Y}) &= P(\overline{X} - \overline{Y} > 0) \\
&= P\left(T > \frac{0 - (43{,}71 - 39{,}63)}{\sqrt{\frac{(9-1)(34{,}5744)+(7-1)(58{,}9824)}{9+7-2}}}\right) \\
&= P\left(T > \frac{-4{,}08}{6{,}71{,}08}\right) \\
&= P(T > -0{,}60796) \\
&= 0{,}72328
\end{aligned}$$

La probabilidad de que la media muestral del grupo mX sea mayor a la media muestral del grupo Y es 0.72328. ■

Caso (b): $\sigma_x^2 \neq \sigma_y^2$:

Sea S_X^2 la varianza muestral de $X_1, X_2, \ldots, X_{\mathbf{m}}$ y sea S_Y^2 la varianza muestral de $Y_1, Y_2, \ldots, Y_{\mathbf{n}}$. Si $\sigma_x^2 \neq \sigma_y^2$ entonces

$$T = \frac{(\overline{X} - \overline{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{S_X^2}{\mathbf{m}} + \frac{\S_Y^2}{\mathbf{n}}}} \qquad (7.1)$$

se puede demostrar que tiene una distribución "t" de Student con algunos grados de libertad, que se denota por ν. Actualmente, no existe una fórmula exacta para los grados de libertad ν en el caso de que ambas varianzas poblacionales sean desconocidas y no se puedan asumir iguales. La denominada ecuación de Welch-Satterthwaite se utiliza para calcular una aproximación a los grados de libertad efectivos de una combinación lineal de varianzas muestrales independientes, también conocidas como grados de libertad agrupados. Esta ecuación está dada por

$$\nu = \frac{\left(\dfrac{S_X^2}{\mathbf{m}} + \dfrac{S_Y^2}{\mathbf{n}}\right)^2}{\dfrac{\left(\dfrac{S_X^2}{\mathbf{m}}\right)^2}{\mathbf{m}-1} + \dfrac{\left(\dfrac{S_Y^2}{\mathbf{n}}\right)^2}{\mathbf{n}-1}}$$

y se usa comúnmente en escenarios en los que se desconocen dos variaciones de población y no hay evidencia para suponer que son iguales.

En consecuencia, en esta situación, se puede construir un intervalo de confianza del $100(1-\alpha)\,\%$ para $\mu_X - \mu_Y$ como

$$I[\mu_X - \mu_Y : 100(1-\alpha)\,\%] = (\overline{X} - \overline{Y}) \pm t(\nu, 1-\alpha/2)\sqrt{\frac{S_X^2}{\mathbf{m}} + \frac{S_Y^2}{\mathbf{n}}}.$$

Ejemplo 7.3. Los siguientes resultados se obtuvieron de un experimento que comparó los tiempos de absorción (en minutos) de dos medicamentos administrados por vía oral. El científico que realiza el experimento tiene fuertes sospechas de que las variaciones de los tiempos de absorción de ambos medicamentos no son iguales.

Druga X :	$\mathbf{m} = 14$	$\sum_x = 301{,}6$	$\sum x^2 = 8120{,}20$
Druga Y :	$\mathbf{n} = 10$	$\sum y = 210{,}0$	$\sum y^2 = 5207{,}49$

Encuentre un intervalo de confianza del 95 % para la diferencia entre los tiempos medios de absorción de los dos medicamentos.

Solución:
El intervalo de confianza del 95 % para la diferencia entre la media Los tiempos de absorción de los dos fármacos están dados por

$$I[\mu_X - \mu_Y : 100(1-\alpha)\,\%] = (\overline{X} - \overline{Y}) \pm t(\nu, 1-\alpha/2)\sqrt{\frac{S_X^2}{\mathbf{m}} + \frac{S_Y^2}{\mathbf{n}}}.$$

Las varianzas muestrales para cada medicamento están dadas por

$$S_X^2 = \frac{1}{\mathbf{m}-1}\left[\sum_{i=1}^{\mathbf{m}} X_i^2 - \frac{1}{\mathbf{m}}\left(\sum_{i=1}^{\mathbf{m}} X_i\right)^2\right] = \frac{1}{13}[8120{,}72 - \frac{1}{14}(3o1{,}6)^2]$$

$$= 124{,}8764835$$

$$S_Y^2 = \frac{1}{\mathbf{n}-1}\left[\sum_{i=1}^{\mathbf{n}} Y_i^2 - \frac{1}{\mathbf{n}}\left(\sum_{i=1}^{\mathbf{n}} Y_i\right)^2\right] = \frac{1}{9}[5207{,}49 - \frac{1}{10}(210{,}0)^2]$$

$$= 88{,}61$$

y la aproximación de Welch-Satterthwaite viene dada por

$$\nu = \frac{\left(\frac{S_X^2}{\mathbf{m}} + \frac{S_Y^2}{\mathbf{n}}\right)^2}{\frac{\left(\frac{S_X^2}{\mathbf{m}}\right)^2}{\mathbf{m}-1} + \frac{\left(\frac{S_Y^2}{\mathbf{n}}\right)^2}{\mathbf{n}-1}} = \frac{\left(\frac{124{,}8764835}{14} + \frac{88{,}61}{10}\right)^2}{\frac{\left(\frac{124{,}8764835}{14}\right)^2}{13} + \frac{\left(\frac{88{,}61}{10}\right)^2}{9}}$$

$$= 21{,}29808397 = 21.$$

Por lo tanto, el intervalo de confianza del 95 % para la diferencia entre los tiempos medios de absorción de los dos fármacos es

$$I[\mu_X - \mu_Y : 100(1-\alpha)\,\%] = (\overline{X} - \overline{Y}) \pm t(\nu, 1-\alpha/2)\sqrt{\frac{S_X^2}{\mathbf{m}} + \frac{S_Y^2}{\mathbf{n}}}.$$

$$= (21{,}54285714 - 21) \pm t(21, 0, 975)\sqrt{\frac{124{,}8764835}{14} + \frac{88{,}61}{10}}$$
$$= 0{,}54285714 \pm 8{,}770782844$$

$$I[\mu_X - \mu_Y : 95\,\%] = [-8{,}2279, 9{,}3136]$$

con cuatro decimales donde $t0{,}975(21) = 2{,}080$. ■