Ingeniería Matemática
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE
MA 3402, 9/10/21, Prof. R. Gouet
Profs. Aux. B. Hernández y S. López

Solución Control # 1

1. Un ingeniero debe medir los 3 ángulos $\theta_1, \theta_2, \theta_3$ de un triángulo, para lo cual realiza n mediciones independientes de cada ángulo, siendo X_{ij} la i-ésima medición del ángulo $\theta_j, i = 1, \dots, n; j = 1, 2, 3$. Suponga que las observaciones X_{ij}, son normales, con esperanzas $\theta_j, j = 1, 2, 3$ y varianza común σ^2 desconocida.

   (a) (1 pt.) Plantee en detalle el modelo paramétrico correspondiente, es decir, espacio de observaciones, espacio de parámetros y familia de probabilidades, para la observación
   $$\mathbf{X} = (X_{11}, \dots, X_{n1}, X_{12}, \dots, X_{n2}, X_{13}, \dots, X_{n3}).$$

   (b) (2 pts.) Obtenga los EMV $\hat{\theta}_j$ de los ángulos θ_j, bajo la condición de que la suma de los ángulos es una constante conocida. Calcule sesgo y varianza de los estimadores.

   (c) (2 pts.) Obtenga los EMV $\tilde{\theta}_j$ de los ángulos θ_j, ignorando la condición sobre los ángulos. Calcule sesgo y varianza de los estimadores y compare con los resultados del apartado (b).

   (d) (2 pts.) Usando el método del pivote proponga un intervalo de confianza más corto, de nivel $1 - \alpha$, para el número π, usando los datos X_{ij}.

   Sol:

   (a) Dado que las mediciones se distribuyen normalmente, es razonable proponer a $\mathbb{R}$ como espacio para una componente y, dado que el vector $\mathbf{X}$ tiene largo $3n$, resulta $\mathcal{X} = \mathbb{R}^{3n}$. Las medidas de probabilidad son Gaussianas en $\mathcal{X}$ con densidad
   $$f_{\boldsymbol{\theta}}(\mathbf{x}) = K\sigma^{-3n} e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n (x_{i1}-\theta_1)^2 + \sum_{i=1}^n (x_{i2}-\theta_2)^2 + \sum_{i=1}^n (x_{i3}-\theta_3)^2\right)},$$
   donde $K > 0$ es una constante apropiada (que no depende de los parámetros). Por otra parte, el espacio de parámetros es
   $$\Theta = \{(\theta_1, \theta_2, \theta_3, \sigma) \,|\, \theta_i \in (0, \pi), i = 1, 2, 3; \theta_1 + \theta_2 + \theta_3 = \pi, \sigma > 0\}.$$

   (b) Para calcular los EMV usamos la log verosimilitud en la cual incorporamos directamente la restricción de la suma de ángulos igual a π. Sea entonces
   $$l = -3n \log \sigma - \frac{1}{2\sigma^2} \left(\sum_{i=1}^n (x_{i1} - \theta_1)^2 + \sum_{i=1}^n (x_{i2} - \theta_2)^2 + \sum_{i=1}^n (x_{i3} - (\pi - \theta_1 - \theta_2))^2 \right).$$
   Derivando c/r a θ_1, θ_2 e igualando a 0, llegamos a las ecuaciones
   $$\bar{x}_1 - \theta_1 - \bar{x}_3 + \pi - \theta_1 - \theta_2 = 0,$$
   $$\bar{x}_2 - \theta_2 - \bar{x}_3 + \pi - \theta_1 - \theta_2 = 0.$$
   Equivalentes a
   $$2\theta_1 + \theta_2 = \bar{x}_1 - \bar{x}_3 + \pi,$$

$$\theta_1 + 2\theta_2 = \bar{x}_2 - \bar{x}_3 + \pi,$$

De donde se obtienen las soluciones

$$\hat{\theta}_1 := \bar{x}_1 - (\bar{x}_1 + \bar{x}_2 + \bar{x}_3 - \pi)/3 = \bar{x}_1 - (\bar{x} - \pi/3),$$

$$\hat{\theta}_2 := \bar{x}_2 - (\bar{x}_1 + \bar{x}_2 + \bar{x}_3 - \pi)/3 = \bar{x}_2 - (\bar{x} - \pi/3),$$

$$\hat{\theta}_3 := \bar{x}_3 - (\bar{x}_1 + \bar{x}_2 + \bar{x}_3 - \pi)/3 = \bar{x}_3 - (\bar{x} - \pi/3),$$

donde $\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$ y $\bar{x} = \frac{1}{3n}\sum_{i,j} x_{ij}$. Aprovechamos de comprobar que se cumple

$$\hat{\theta}_1 + \hat{\theta}_2 + \hat{\theta}_3 = \pi.$$

Para ver el sesgo basta analizar uno de los estimadores. Por simetría, las conclusiones de uno valen para los demás.

$$\mathbb{E}_{\boldsymbol{\theta}}(\hat{\theta}_1(\mathbf{X})) = \mathbb{E}_{\boldsymbol{\theta}}(\bar{X}_1 - (\bar{X}_1 + \bar{X}_2 + \bar{X}_3 - \pi)/3) = \theta_1 - (\theta_1 + \theta_2 + \theta_3 - \pi)/3 = \theta_1.$$

Concluimos que los EMV sujetos a la restricción son insesgados. Ahora calculamos las varianzas, teniendo presente que las $\bar{X}_j$ son independientes y tienen varianzas σ^2/n.

$$Var_{\boldsymbol{\theta}}(\hat{\theta}_1(\mathbf{X})) = Var_{\boldsymbol{\theta}}(\bar{X}_1 - (\bar{X}_1 + \bar{X}_2 + \bar{X}_3 - \pi)/3) = (\frac{4}{9} + \frac{2}{9})\frac{\sigma^2}{n} = \frac{2\sigma^2}{3n}$$

(c) Calculamos los EMV sin restricción.

$$l = -3n \log \sigma - \frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_{i1} - \theta_1)^2 + \sum_{i=1}^{n}(x_{i2} - \theta_2)^2 + \sum_{i=1}^{n}(x_{i3} - \theta_3)^2\right).$$

Derivando c/r a $\theta_j, j = 1, 2, 3$ e igualando a 0 llegamos a las ecuaciones

$$\sum_{i=1}^{n}(x_{ij} - \theta_j) = 0,$$

cuyas soluciones son $\tilde{\theta}_j := \bar{x}_j, j = 1, 2, 3$. Se comprueba fácilmente que los $\tilde{\theta}_j$ son minimizantes y corresponden justamente a los EMV de los θ_j, sin tomar en cuenta la restricción de los ángulos. El sesgo y la varianza de los estimadores se obtienen directamente de las propiedades de la media de variables aleatorias gaussianas. Resulta entonces que $\mathbb{E}_{\boldsymbol{\theta}}(\bar{X}_j) = \theta_j$ y $Var_{\boldsymbol{\theta}}(\bar{X}_j) = \sigma^2/n, j = 1, 2, 3$.

De la comparación con los resultados del apartado (b) concluimos que ambos estimadores (con y sin restricción) son insesgados pero los estimadores que ignoran la restricción tienen mayor varianza (50% más).

(d) Notemos que $\bar{X}_1 + \bar{X}_2 + \bar{X}_3$ es una variable gaussiana con esperanza $\theta_1 + \theta_2 + \theta_3 = \pi$ y varianza $3\frac{\sigma^2}{n}$ Por lo tanto,

$$Z = \sqrt{\frac{n}{3}}\frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3 - \pi}{\sigma} = \sqrt{\frac{n}{3}}\frac{3\bar{X} - \pi}{\sigma}$$

tiene ley normal $N(0, 1)$. Si σ fuera conocido podríamos usar Z como pivote para π y obtener

$$I(\mathbf{x}) = \left[3\bar{x} - \frac{\sigma}{\sqrt{n/3}}\Phi^{-1}\Big(1 - \frac{\alpha}{2}\Big)\, ,\, 3\bar{x} + \frac{\sigma}{\sqrt{n/3}}\Phi^{-1}\Big(1 - \frac{\alpha}{2}\Big)\right].$$

Pero, dado que σ no se conoce, podemos estimarlo con, por ejemplo $\hat{\sigma}_j^2 = \frac{1}{n-1}\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$, con lo cual obtendríamos el intervalo

$$I_j(\mathbf{x}) = \left[3\bar{x} - \frac{\hat{\sigma}_j}{\sqrt{n/3}}\Phi_{n-1}^{-1}\Big(1 - \frac{\alpha}{2}\Big)\,,\, 3\bar{x} + \frac{\hat{\sigma}_j}{\sqrt{n/3}}\Phi_{n-1}^{-1}\Big(1 - \frac{\alpha}{2}\Big)\right].$$

También es posible mejorar la estimación de σ usando la totalidad de la información, definiendo $\hat{\sigma}^2 := (\hat{\sigma}_1^2 + \hat{\sigma}_2^2 + \hat{\sigma}_3^2)/3$. En este caso se mejora la calidad de la estimación de σ y la t de student tiene ahora $3n - 3$ grados de libertad. El intervalo resultante, que es más corto que el anterior debido al aumento de los grados de libertad, está dado por

$$\hat{I}(\mathbf{x}) = \left[3\bar{x} - \frac{\hat{\sigma}}{\sqrt{n/3}}\Phi_{3n-3}^{-1}\Big(1 - \frac{\alpha}{2}\Big)\,,\, 3\bar{x} + \frac{\hat{\sigma}}{\sqrt{n/3}}\Phi_{3n-3}^{-1}\Big(1 - \frac{\alpha}{2}\Big)\right].$$

2. Se desea determinar lo más precisamente posible el área de un triángulo de altura h y base b, ambas magnitudes desconocidas. Para ello se realizan n mediciones independientes de la altura $\mathbf{X} = (X_1, \ldots, X_n)$ y n mediciones independientes de la base $\mathbf{Y} = (Y_1, \ldots, Y_n)$. Se sabe que las mediciones están sujetas a errores que pueden describirse como variables aleatorias normales independientes $N(0, \sigma^2)$, con σ conocido. Es decir, las observaciones pueden representarse como $X_i = h + \epsilon_i$, $Y_i = b + \delta_i$, $i = 1, \ldots, n$, donde los ϵ_i, δ_i son va iid $N(0, \sigma^2)$.

   (a) (2 pts.) Muestre que el estadístico $(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i)$ es suficiente completo para $\theta = (h, b)$.

   (b) (1 pt.) Muestre que $X_1Y_1/2$ es un estimador insesgado del área del triángulo.

   (c) (2 pts.) Pruebe que $\overline{X}\,\overline{Y}/2$ es EIVUM del área del triángulo.

   (d) (2 pts.) Compruebe que $\frac{1}{2n}\sum_{i=1}^n X_iY_i$ es estimador insesgado del área y explique en qué sentido es peor que $\overline{X}\,\overline{Y}/2$ pero mejor que $X_1Y_1/2$.

   Sol:

   (a) La densidad conjunta de las observaciones está dada por

$$f_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{y}) = K\sigma^{-2n} e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n (x_i - h)^2 + \sum_{i=1}^n (y_i - b)^2\right)}.$$

   Notemos que $f_{\boldsymbol{\theta}}$ pertenece a la familia exponencial y puede ser escrita como

$$f_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{y}) = K e^{\frac{h}{\sigma^2}\sum_{i=1}^n x_i + \frac{b}{\sigma^2}\sum_{i=1}^n y_i - A(h,b)} H(\mathbf{x}, \mathbf{y}).$$

   Observamos que es de rango completo y se concluye de inmediato, por propiedad general de la clase exponencial (Prop. 2.43) que $T(\mathbf{x}, \mathbf{y}) = (\sum_{i=1}^n y_i, \sum_{i=1}^n y_i)$ es suficiente minimal. La completitud es consecuencia de la Prop. 2.49.

   (b) $\mathbb{E}_{\boldsymbol{\theta}}(X_1Y_1/2) = \mathbb{E}_{\boldsymbol{\theta}}(X_1)\mathbb{E}_{\boldsymbol{\theta}}(Y_1)/2 = hb/2$, es decir, el estimador $X_1Y_1/2$ es insesgado.

   (c) Para comprobar que $\overline{X}\,\overline{Y}/2$ es EIVUM del área del triángulo comprobamos primero que es insesgado. En efecto $\mathbb{E}_{\boldsymbol{\theta}}(\bar{X}\bar{Y}/2) = \mathbb{E}_{\boldsymbol{\theta}}(\bar{X})\mathbb{E}_{\boldsymbol{\theta}}(\bar{Y})/2 = hb/2$, debido a que las medias son estimadores insesgados de las esperanzas. Por otra parte, $\bar{X}\bar{Y}/2$ es función del estadístico suficiente y completo $T(\mathbf{x}, \mathbf{y})$, definido en el apartado (a). Luego, gracias al Teorema 2.56, de Lehmann y Scheffé, concluimos que es EIVUM.

   (d) Del apartado (b) sabemos que $X_iY_i/2$ es estimador insesgado de $hb/2$, por lo tanto, la media aritmética de los $X_iY_i/2$, también lo es.

Notamos que este estimador no es función el estadístico suficiente completo. Vamos a comprobar que su varianza es mayor que la del EIVUM. Calculamos primero

$$Var_{\boldsymbol{\theta}}\left(\frac{1}{2n}\sum_{i=1}^{n} X_iY_i\right) = \frac{1}{4n^2}\sum_{i=1}^{n} Var_{\boldsymbol{\theta}}(X_iY_i) = \frac{1}{4n}Var_{\boldsymbol{\theta}}(X_1Y_1).$$

Por otra parte,

$$\begin{aligned}
Var_{\boldsymbol{\theta}}(X_1Y_1) &= \mathbb{E}_{\boldsymbol{\theta}}(X_1^2Y_1^2) - (\mathbb{E}_{\boldsymbol{\theta}}(X_1Y_1))^2 \\
&= \mathbb{E}_{\boldsymbol{\theta}}(X_1^2)\mathbb{E}_{\boldsymbol{\theta}}(Y_1^2) - (\mathbb{E}_{\boldsymbol{\theta}}(X_1)\mathbb{E}_{\boldsymbol{\theta}}(Y_1))^2 \\
&= (\sigma^2 + h^2)(\sigma^2 + b^2) - h^2b^2 \\
&= \sigma^4 + \sigma^2(h^2 + b^2).
\end{aligned}$$

De manera que

$$Var_{\boldsymbol{\theta}}\left(\frac{1}{2n}\sum_{i=1}^{n} X_iY_i\right) = \frac{\sigma^2}{4n}(\sigma^2 + h^2 + b^2).$$

De manera análoga obtenemos

$$\begin{aligned}
Var_{\boldsymbol{\theta}}(\bar{X}\bar{Y}) &= \mathbb{E}_{\boldsymbol{\theta}}(\bar{X}^2\bar{Y}^2) - (\mathbb{E}_{\boldsymbol{\theta}}(\bar{X}\bar{Y}))^2 \\
&= \mathbb{E}_{\boldsymbol{\theta}}(\bar{X}^2)\mathbb{E}_{\boldsymbol{\theta}}(\bar{Y}^2) - (\mathbb{E}_{\boldsymbol{\theta}}(\bar{X})\mathbb{E}_{\boldsymbol{\theta}}(\bar{Y}))^2 \\
&= (\sigma^2/n + h^2)(\sigma^2/n + b^2) - h^2b^2 \\
&= \sigma^4/n^2 + (\sigma^2/n)(h^2 + b^2).
\end{aligned}$$

Por lo tanto, la varianza del EIVUM es

$$Var_{\boldsymbol{\theta}}(\bar{X}\bar{Y}/2) = \frac{\sigma^2}{4n}\left(\frac{\sigma^2}{n} + h^2 + b^2\right).$$

Terminamos con el cociente de las varianzas:

$$\frac{Var_{\boldsymbol{\theta}}(\bar{X}\bar{Y}/2)}{Var_{\boldsymbol{\theta}}\left(\frac{1}{2n}\sum_{i=1}^{n} X_iY_i\right)} = \frac{\frac{\sigma^2}{4n}\left(\frac{\sigma^2}{n} + h^2 + b^2\right)}{\frac{\sigma^2}{4n}(\sigma^2 + h^2 + b^2)} = \frac{\frac{\sigma^2}{n} + h^2 + b^2}{\sigma^2 + h^2 + b^2} \leq 1,$$

para concluir que el EIVUM es mejor, como era esperable. La varianza de $X_1Y_1/2$ es obviamente mayor que la del promedio aritmético.

3. Suponga que $X_1, \ldots, X_n$ son va independientes pero no idénticamente distribuidas, tales que X_i es normal $N(\mu a_i, 1), i = 1, \ldots, n$, donde los a_i son constantes conocidas.

   (a) (3 pts.) Desarrolle el TNP de nivel $\alpha_0 \in (0, 1)$, para $H_0 : \mu = \mu_0$ vs $H_1 : \mu = \mu_1$, donde $\mu_0 < \mu_1$ son conocidos. Obtenga la región crítica en términos de las constantes del problema $(\alpha_0, a_i, \mu_0, \mu_1)$ y de la función Φ.

   (b) (2 pts.) Calcule la potencia del TNP del apartado anterior, dejando todo expresado en términos de las constantes del problema y de Φ.

   (c) (1 pts.) Suponga que usted puede escoger las constantes $a_i, i = 1, \ldots, n$. Muestre que para maximizar la potencia calculada en (b) conviene escoger todos los a_i iguales. Indique cómo se simplifica el TNP en tal situación.

Sol:

(a) La densidad de X_i es $\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-a_i\mu)^2}$, por lo tanto, la densidad de $\mathbf{X}$ está dada por

$$f_\mu(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^n (x_i - a_i\mu)^2}.$$

Para desarrollar el TNP debemos simplificar la desigualdad $f_{\mu_1} \geq k f_{\mu_0}$ vía equivalencias.

$$\frac{f_{\mu_1}(\mathbf{x})}{f_{\mu_0}(\mathbf{x})} = e^{-\frac{1}{2}(\sum_{i=1}^n (x_i - a_i\mu_1)^2 - \sum_{i=1}^n (x_i - a_i\mu_0)^2)} \geq k \Leftrightarrow \sum_{i=1}^n a_i x_i \geq k'.$$

Ahora imponemos el nivel del test, resolviendo

$$\mathbb{P}_{\mu_0}\left(\sum_{i=1}^n a_i X_i \geq k'\right) = \alpha_0,$$

equivalente a

$$\mathbb{P}_{\mu_0}\left(\frac{\sum_{i=1}^n a_i X_i - \mu_0||a||^2}{||a||} \geq \frac{k' - \mu_0||a||^2}{||a||}\right) = \alpha_0,$$

donde $||a|| = \sqrt{\sum_{i=1}^n a_i^2}$. O bien,

$$1 - \Phi\left(\frac{k' - \mu_0||a||^2}{||a||}\right) = \alpha_0,$$

de donde obtenemos

$$k' = \mu_0||a||^2 + ||a||\Phi^{-1}(1-\alpha_0).$$

(b) Para la potencia calculamos

$$\mathbb{P}_{\mu_1}\left(\sum_{i=1}^n a_i X_i \geq \mu_0||a||^2 + ||a||\Phi^{-1}(1-\alpha_0)\right),$$

equivalente a

$$\begin{aligned}\pi(\mu_1) &= \mathbb{P}_{\mu_1}\left(\frac{\sum_{i=1}^n a_i X_i - \mu_1||a||^2}{||a||} \geq \Phi^{-1}(1-\alpha_0) - \frac{(\mu_1-\mu_0)||a||^2}{||a||}\right) \\ &= 1 - \Phi\left(\Phi^{-1}(1-\alpha_0) - (\mu_1-\mu_0)||a||\right).\end{aligned}$$

Notamos que, debido a que $\mu_0 < \mu_1$, $\pi(\mu_1)$ es función creciente de $||a||$, que se maximiza cuando $||a|| = \infty$. Suponiendo que los a_i son finitos en el sentido que $||a|| \leq C$, nada indica que sirva que los a_i sean iguales.