Ingeniería Matemática
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE
MA 3402, 13/11/21, Prof. R. Gouet
Profs. Aux. B. Hernández y S. López

Solución Control # 2

1. Sean $X_1, X_2, \ldots, X_n$ va iid con ley normal $N(\mu, 1)$. El objetivo del ejercicio es desarrollar el TRV de nivel α para $H_0 : \mu = 0 \quad \text{vs} \quad H_1 : |\mu| \geq \mu_1 > 0$, donde μ_1 es conocido.

   (a) (1 pt.) Especifique $\mathcal{X}$, $\Theta, \Theta_0, \Theta_1$

   Solución: Puesto que trabajamos con una muestra de va normales, de tamaño n, el espacio de las observaciones es $\mathcal{X} = \mathbb{R}^n$. El parámetro desconocido es μ, que puede variar, en principio, en todo $\mathbb{R}$. Sin embargo, considerando las hipótesis $H_0 : \mu = 0$ y $H_1 : |\mu| \geq \mu_1 > 0$, resulta que $\Theta_0 = \{0\}, \Theta_1 = (-\infty, -\mu_1] \cup [\mu_1, \infty)$ y $\Theta = \Theta_0 \cup \Theta_1$.

   (b) (2 pts.) Compruebe que $\lambda(\mathbf{x})$ del TRV puede escribirse como

   $$\lambda(\mathbf{x}) = e^{-n\mu_1^2/2} e^{n\mu_1 t} \mathbf{1}_{\{t \leq \mu_1\}} + e^{nt^2/2} \mathbf{1}_{\{t > \mu_1\}},$$

   donde $t = |\bar{x}|$. Indicación: para calcular el supremo en Θ_1 se sugiere considerar los casos $|\bar{x}| > \mu_1$, $0 \leq \bar{x} \leq \mu_1$ y $-\mu_1 \leq \bar{x} < 0$.

   Solución: Para calcular $\lambda(\mathbf{x})$ se requiere tener los supremos sobre Θ_0 y Θ_1. Dado que $\Theta_0 = \{0\}$, basta tomar la verosimilitud con $\mu = 0$. Es decir, siendo

   $$L(\mu, \mathbf{x}) = K e^{-\frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2}$$

   la función de verosimilitud (en general), se tiene

   $$\sup_{\mu \in \Theta_0} L(\mu, \mathbf{x}) = L(0, \mathbf{x}) = K e^{-\frac{1}{2} \sum_{i=1}^n x_i^2},$$

   donde K es una constante apropiada. Por otra parte, para el supremo en Θ_1 consideramos los casos siguientes:

   i. $|\bar{x}| > \mu_1$. Sabiendo que el máximo de la verosimilitud se alcanza en $\mu = \bar{x}$ y que $\bar{x}$ está en Θ_1, entonces

   $$\sup_{\mu \in \Theta_1} L(\mu, \mathbf{x}) = L(\bar{x}, \mathbf{x}) = K e^{-\frac{1}{2} \sum_{i=1}^n (x_i - \bar{x})^2}$$

   y

   $$\lambda(\mathbf{x}) = \frac{\sup_{\mu \in \Theta_1} L(\mu, \mathbf{x})}{\sup_{\mu \in \Theta_0} L(\mu, \mathbf{x})} = e^{\frac{n}{2}\bar{x}^2} = e^{\frac{n}{2}t^2},$$

   donde $t = |\bar{x}|$.

   ii. $0 \leq \bar{x} \leq \mu_1$. En este caso $\bar{x}$ está del lado positivo y, dada la forma de la verosimilitud (como función de μ), para maximizar conviene acercarse lo más posible hacia $\bar{x}$, de manera que el óptimo se encuentra en μ_1. En tal caso (notando que aquí $t = \bar{x}$)

   $$\sup_{\mu \in \Theta_1} L(\mu, \mathbf{x}) = L(\mu_1, \mathbf{x}) = K e^{-\frac{1}{2} \sum_{i=1}^n (x_i - \mu_1)^2}$$

   y

   $$\lambda(\mathbf{x}) = \frac{\sup_{\mu \in \Theta_1} L(\mu, \mathbf{x})}{\sup_{\mu \in \Theta_0} L(\mu, \mathbf{x})} = e^{-\frac{1}{2}\left(\sum_{i=1}^n (x_i - \mu_1)^2 - \sum_{i=1}^n x_i^2\right)} = e^{\frac{n}{2}\mu_1(2\bar{x} - \mu_1)} = e^{\frac{n}{2}\mu_1(2t - \mu_1)}.$$

iii. $-\mu_1 \leq \bar{x} < 0$. Ahora $\bar{x}$ está del lado negativo, de manera que el óptimo lo encontramos en $-\mu_1$ y tenemos
$$\sup_{\mu \in \Theta_1} L(\mu, \mathbf{x}) = L(-\mu_1, \mathbf{x}) = Ke^{-\frac{1}{2}\sum_{i=1}^n (x_i+\mu_1)^2}.$$
Entonces,
$$\lambda(\mathbf{x}) = \frac{\sup_{\mu \in \Theta_1} L(\mu, \mathbf{x})}{\sup_{\mu \in \Theta_0} L(\mu, \mathbf{x})} = e^{-\frac{1}{2}\left(\sum_{i=1}^n (x_i+\mu_1)^2 - \sum_{i=1}^n x_i^2\right)} = e^{\frac{n}{2}\mu_1(-2\bar{x}-\mu_1)} = e^{\frac{n}{2}\mu_1(2t-\mu_1)}.$$
Los cálculos anteriores muestran que $\lambda(\mathbf{x})$ tiene la forma
$$\lambda(\mathbf{x}) = e^{-n\mu_1^2/2} e^{n\mu_1 t} \mathbf{1}_{\{t \leq \mu_1\}} + e^{nt^2/2} \mathbf{1}_{\{t > \mu_1\}}.$$

(c) (2 pts.) Muestre que el TRV tiene región crítica de la forma
$$R = \{x \in \mathcal{X} \mid |\bar{x}| \geq k_\alpha\}.$$
Solución: De la parte b) podemos concluir fácilmente que $\lambda(\mathbf{x})$ es función creciente de t, por lo tanto, la región crítica tiene la forma indicada, puesto que $\lambda(\mathbf{x}) \geq k \Leftrightarrow |\bar{x}| \geq k'$.

(d) (1 pt.) Determine k_α en términos de Φ (función de distribución de una va $N(0,1)$).
Solución: Para calcular k_α usamos que, bajo H_0, $\bar{X}$ se distribuye como $N(0, 1/n)$. Resolvemos
$$P[|\bar{X}| \geq k_\alpha] = P[\sqrt{n}|\bar{X}| \geq \sqrt{n}k_\alpha] = 2(1 - \Phi(\sqrt{n}k_\alpha)) = \alpha,$$
obteniendo
$$k_\alpha = \frac{1}{\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right),$$
donde Φ denota la función de distribución de una va $N(0,1)$ y Φ^{-1} su inversa.

2. Sean $X_1, X_2, \ldots, X_n$ v.a. independientes, normales con varianza común σ^2 conocida. Suponga que $E(X_1) = \mu$ y $E(X_i) = \lambda E(X_{i-1}), i = 2, \ldots, n$, donde $\mu \in \mathbb{R}$ es el parámetro desconocido y $\lambda > 0$ es una constante conocida.

(a) (3 pts.) Considere como densidad a priori $\pi(\mu)$ la normal $N(0,1)$. Calcule la densidad a posteriori de μ y los estimadores de Bayes para μ y μ^2.
Solución: Las X_i tienen ley normal $N(\mu\lambda^{i-1}, 1)$ de donde obtenemos (siendo K constante positiva que puede cambiar en las distintas líneas y sirve para que la densidad integre a 1)
$$f(\mathbf{x} \mid \mu) = Ke^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu\lambda^{i-1})^2}$$
y $\pi(\mu) = Ke^{-\frac{\mu^2}{2}}$. Por lo tanto,
$$\begin{aligned}
f(\mu \mid \mathbf{x}) &= Ke^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu\lambda^{i-1})^2 - \frac{\mu^2}{2}} \\
&= Ke^{-\frac{1}{2\sigma^2}\left(\mu^2\left(\sigma^2 + \sum_{i=1}^n \lambda^{2i-2}\right) - 2\mu\sum_{i=1}^n x_i\lambda^{i-1}\right)} \\
&= Ke^{-\frac{\sigma^2 + \sum_{i=1}^n \lambda^{2i-2}}{2\sigma^2}\left(\mu^2 - 2\mu\frac{\sum_{i=1}^n x_i\lambda^{i-1}}{\sigma^2 + \sum_{i=1}^n \lambda^{2i-2}}\right)} \\
&= Ke^{-\frac{1}{2\tau^2}\left(\mu^2 - 2\mu\nu + \nu^2\right)} \\
&= Ke^{-\frac{1}{2\tau^2}(\mu - \nu)^2},
\end{aligned}$$

donde $\tau^2 = \frac{\sigma^2}{\sigma^2+\sum_{i=1}^n \lambda^{2i-2}}$ y $\nu = \frac{\sum_{i=1}^n x_i\lambda^{i-1}}{\sigma^2+\sum_{i=1}^n \lambda^{2i-2}}$. Por lo tanto, la densidad a posteriori $f(\mu\,|\,\mathbf{x})$ es normal $N(\nu,\tau^2)$.

Respecto del estimador de Bayes sabemos que es la esperanza a posteriori. Así resulta que $\hat{\mu}_B = \mathbb{E}(\mu\,|\,\mathbf{x}) = \nu$ y el estimador de Bayes de μ^2 es $\mathbb{E}(\mu^2\,|\,\mathbf{x})$. Para calcular esta última esperanza notemos que una v.a. $N(\nu,\tau^2)$ tiene la misma distribución que $\tau Z+\nu$, donde Z es $N(0,1)$. Por lo tanto

$$\mathbb{E}(\mu^2\,|\,\mathbf{x}) = \mathbb{E}((\tau Z+\nu)^2) = \mathbb{E}(\tau^2 Z^2 + 2\nu\tau Z + \nu^2) = \tau^2+\nu^2.$$

(b) (3 pts.) Suponga ahora que π es la densidad a priori impropia, dada por $\pi(\mu)=1, \mu\in\mathbb{R}$. Muestre que existe y calcule la densidad a posteriori $f(\mu\,|\,\mathbf{x})$. Obtenga los estimadores MAP y de Bayes para μ.

Solución: En este caso trabajamos como en el apartado anterior pero con $\pi(\mu)=1$. Tenemos entonces

$$\begin{aligned}
f(\mu\,|\,\mathbf{x}) &= Ke^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu\lambda^{i-1})^2}\\
&= Ke^{-\frac{1}{2\sigma^2}\left(\mu^2\sum_{i=1}^n \lambda^{2i-2} - 2\mu\sum_{i=1}^n x_i\lambda^{i-1}\right)}\\
&= Ke^{-\frac{\sum_{i=1}^n \lambda^{2i-2}}{2\sigma^2}\left(\mu^2-2\mu\frac{\sum_{i=1}^n x_i\lambda^{i-1}}{\sum_{i=1}^n \lambda^{2i-2}}\right)}\\
&= Ke^{-\frac{1}{2\tau^2}\left(\mu^2-2\mu\nu+\nu^2\right)}\\
&= Ke^{-\frac{1}{2\tau^2}(\mu-\nu)^2},
\end{aligned}$$

donde

$$\tau^2 = \frac{\sigma^2}{\sum_{i=1}^n \lambda^{2i-2}}, \qquad \nu = \frac{\sum_{i=1}^n x_i\lambda^{i-1}}{\sum_{i=1}^n \lambda^{2i-2}}.$$

De lo anterior se desprende que $f(\mu\,|\,\mathbf{x})$ es $N(\nu,\tau^2)$ y, por lo tanto, dado que la densidad normal alcanza su máximo en la esperanza, resulta que $\hat{\mu}_B = \hat{\mu}_{MAP} = \nu$.

3. Una persona desea comprar 1 Kg. de queso en láminas, c/u de las cuales pesa 50 grs. Debido al proceso de fabricación, algunas de estas láminas (θ) tienen buen sabor y otras no. Un lámina de buen sabor sirve para un sándwich que al venderse produce una utilidad p y una de mal sabor producirá una pérdida q, por la necesidad de indemnizar al cliente afectado y por el queso perdido. Para decidir su compra, la persona escoge una lámina al azar y la prueba. El resultado del test es el sabor, bueno o malo, de la lámina que probó. Ocurre que para probar una lámina, la persona debe cancelar su valor, que asciende a r. El problema consiste en decidir de manera racional si compra o no el Kg. de queso.

   (a) (1 pt.) Detalle los elementos que caracterizan este problema de decisión: estados de la naturaleza, conjunto de decisiones, v.a. observada, función de pérdida, etc.

   Solución: El conjunto de estados de la naturaleza es $\Theta=\{0,1,\ldots,20\}$, donde $\theta\in\Theta$ representa el número de láminas sabrosas en el Kg. de queso. El conjunto de decisiones es $D=\{0,1\}$, donde 0 representa la decisión de no comprar y 1 la de comprar el Kg. de queso. La v.a. observada es el estado X de la lámina probada. La v.a. X toma valores en $\mathcal{X}=\{0,1\}$, donde 1 indica que la lámina probada era buena y 0, lo contrario. La función de pérdida está dada por

$$L(d,\theta) = r + d(q(20-\theta)-p\theta).$$

   (b) (1 pt.) Escriba las posibles reglas de decisión en este problema.

Solución: Las posibles reglas de decisión son todas las funciones entre $\mathcal{X}$ y D, que son 4; $\delta_1(x) = 0, \forall x$, equivalente a no comprar, cualquiera sea el resultado de la prueba; $\delta_2(x) = 1, \forall x$, equivalente a comprar, cualquiera sea el resultado de la prueba; $\delta_3(0) = 0, \delta_3(1) = 1$, equivalente a comprar si y sólo si la prueba es satisfactoria y finalmente, $\delta_4(0) = 1, \delta_4(1) = 0$, equivalente a comprar si y sólo si la prueba es insatisfactoria.

(c) (2 pts.) Suponga que $r = 2, p = 4, q = 6$. Calcule los riesgos asociados a todas las reglas de decisión y determine reglas admisibles e inadmisibles.

Solución: Con los valores dados, tenemos

$$L(d, \theta) = 2 + d(6(20 - \theta) - 4\theta).$$

El riesgo de una regla δ se define como $R_\theta(\delta) = \mathbb{E}_\theta(L(\delta(X), \theta))$. Para las reglas definidas previamente tenemos,

$$R_\theta(\delta_1) = \mathbb{E}_\theta(L(\delta_1(X), \theta)) = \mathbb{E}_\theta(L(0, \theta)) = r = 2,$$

$$R_\theta(\delta_2) = \mathbb{E}_\theta(L(\delta_2(X), \theta)) = \mathbb{E}_\theta(L(1, \theta)) = 2 + (6(20 - \theta) - 4\theta) = 122 - 10\theta,$$

$$R_\theta(\delta_3) = L(0, \theta)(1 - \frac{\theta}{20}) + L(1, \theta)\frac{\theta}{20} = 2(1 - \frac{\theta}{20}) + (2 + 6(20 - \theta) - 4\theta)\frac{\theta}{20}.$$

Agrupando términos, obtenemos,

$$R_\theta(\delta_3) = 2 + 6\theta - \theta^2/2.$$

Finalmente,

$$R_\theta(\delta_4) = L(1, \theta)(1 - \frac{\theta}{20}) + L(0, \theta)\frac{\theta}{20} = (2 + 6(20 - \theta) - 4\theta)(1 - \frac{\theta}{20}) + 2\frac{\theta}{20}.$$

Agrupando términos, obtenemos,

$$R_\theta(\delta_4) = 122 - 16\theta + \theta^2/2.$$

Una regla δ es admisible si no existe ninguna otra regla que la domine, en el sentido de tener riesgo uniformemente menor. Podemos graficar las funciones de riesgo para constatar que todas las reglas son admisibles. Notamos que todas las funciones de riesgo se cruzan en el punto de coordenadas $(12, 2)$ y podría parecer que δ_4 es inadmisible, sin embargo, los riesgos de δ_4 y δ_3 se cruzan además en el punto de coordenadas $(10, 12)$, lo cual significa que δ_4 es mejor que δ_3 en el intervalo $]10, 12[$, cuya intersección con Θ es $\{11\}$. Justamente en este punto tenemos $R_{11}(\delta_3) = 7.5 > R_{11}(\delta_4) = 6.5$.

(d) (2 pts.) Considere en (iii) una probabilidad a priori uniforme, en el conjunto de valores posibles de θ y determine la correspondiente solución bayesiana del problema.

Solución: Sea

$$\bar{R}(\delta) = \sum_{\theta=0}^{20} R_\theta(\delta)\pi(\theta) = \frac{1}{21}\sum_{\theta=0}^{20} R_\theta(\delta)$$

el riesgo bayesiano de la regla δ. Para δ_1 tenemos evidentemente, $\bar{R}(\delta_1) = 2$. Por otra parte, $\bar{R}(\delta_2) = 22$, $\bar{R}_{\delta_3} = -19/3$ y $\bar{R}(\delta_4) = 91/3$. Concluimos que la regla de Bayes es δ_3, lo cual es intuitivamente aceptable.