



Ingeniería Eléctrica

FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Análisis de Sistemas Dinámicos y Estimación

EL3204-1

Auxiliar 12 - Estimación

Prof Marcos Orchard - Sebastian Espinoza.
Prof Auxiliar Erik Sáez Aravena.

$\mathcal{I}(\theta)$

Resumen

Este auxiliar cubre conceptos fundamentales de la teoría de estimación de parámetros, que complementa los temas de detección vistos anteriormente. En problemas de detección, el espacio paramétrico Θ era discreto y finito, pero ahora trabajamos con un conjunto **infinito no numerable** de posibles valores para el parámetro que queremos estimar (un continuo).

1. Conceptos Básicos de Estimación

Algunos conceptos importante a considerar son:

1. **Espacio de observación (\mathbb{X}):** Espacio donde la variable aleatoria $X \in \mathbb{X}$ toma valores. Para vectores aleatorios, \mathbb{X}^n representa el espacio producto.
2. **Espacio de decisión (Θ):** En estimación, es un conjunto **infinito no numerable** de valores (a diferencia de detección donde era finito). Ejemplos: $\Theta = \mathbb{R}^+$ para estimar amplitudes, $\Theta = \mathbb{R}$ para estimar medias.
3. **Familia de distribuciones paramétricas (J_θ):** Conjunto de distribuciones indexadas por θ :

$$J_\theta = \{P_X(x|\theta) : \theta \in \Theta\}$$

4. **Estimador:** Una función $\hat{\theta}(x_1, \dots, x_n)$ que depende de las observaciones y entrega una estimación del parámetro θ asociado a la distribución $P_X(x|\theta)$.

2. Propiedades de los Estimadores

1. **Estimador Insesgado:** Un estimador $\hat{\theta}(x_1, \dots, x_n)$ es insesgado si:

$$\mathbb{E}[\hat{\theta}(X_1, \dots, X_n)] = \theta$$

Es decir, en promedio el estimador entrega el valor verdadero del parámetro.

2. **Estimador Asintóticamente Insesgado:** Si el sesgo tiende a cero cuando $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} \mathbb{E}[\hat{\theta}(X_1, \dots, X_n)] = \theta$$

3. **Estimador Consistente:** Un estimador es consistente si converge en probabilidad al valor verdadero:

$$\lim_{n \rightarrow \infty} P_X(|\hat{\theta}(X_1, \dots, X_n) - \theta| > \epsilon) = 0 \quad \forall \epsilon > 0$$

4. **Observación:** Mediante la desigualdad de Chebyshev, si un estimador es insesgado y su varianza tiende a cero cuando $n \rightarrow \infty$, entonces es consistente:

$$\lim_{n \rightarrow \infty} \text{Var}(\hat{\theta}(X_1, \dots, X_n)) = 0 \implies \text{consistencia}$$

3. Condiciones de Regularidad

Las condiciones de regularidad permiten intercambiar el orden de derivación e integración en la función de verosimilitud. Bajo estas condiciones se cumple:

1. **Primera condición:** La esperanza de la derivada de la log-verosimilitud es cero:

$$\mathbb{E}_{X_1^n} \left[\frac{\partial \ln L(X_1, \dots, X_n | \theta)}{\partial \theta} \right] = 0$$

2. **Segunda condición:** La esperanza de un término específico es cero:

$$\mathbb{E}_{X_1^n} \left[\frac{1}{L(X_1, \dots, X_n | \theta)} \frac{\partial^2 L(X_1, \dots, X_n | \theta)}{\partial \theta^2} \right] = 0$$

3. **Identidad de Bartlett:** Relaciona la varianza con su segunda derivada:

$$\mathbb{E}_{X_1^n} \left[\left(\frac{\partial \ln L}{\partial \theta} \right)^2 \right] = -\mathbb{E}_{X_1^n} \left[\frac{\partial^2 \ln L}{\partial \theta^2} \right]$$

Esta identidad permite calcular la información de Fisher de dos formas equivalentes.

4. Información de Fisher

La información de Fisher $\mathcal{I}(\theta)$ cuantifica cuánta información sobre el parámetro θ contienen las observaciones:

$$\mathcal{I}(\theta) = \mathbb{E} \left[\left(\frac{\partial \ln f(X | \theta)}{\partial \theta} \right)^2 \right] = -\mathbb{E} \left[\frac{\partial^2 \ln f(X | \theta)}{\partial \theta^2} \right]$$

Propiedad Aditiva: Para muestras i.i.d., la información de Fisher es aditiva:

$$\mathcal{I}_n(\theta) = n \cdot \mathcal{I}_1(\theta)$$

Esta propiedad refleja que más observaciones independientes proporcionan más información sobre el parámetro.

5. Cota de Cramér-Rao

La cota de Cramér-Rao establece un límite inferior para la varianza de cualquier estimador insesgado:

$$\text{Var}(\hat{\theta}) \geq \frac{1}{\mathcal{I}(\theta)}$$

1. **Estimador Eficiente:** Un estimador insesgado que alcanza la cota de Cramér-Rao se llama **eficiente** o **de mínima varianza**
2. **Interpretación:** La información de Fisher es inversamente proporcional a la mínima varianza posible. Más información \implies menor varianza mínima.

6. Estimador de Máxima Verosimilitud (ML)

El estimador de máxima verosimilitud se obtiene maximizando la función de verosimilitud:

$$\hat{\theta}_{ML} = \arg \max_{\theta} L(x_1, \dots, x_n | \theta) = \arg \max_{\theta} \ln L(x_1, \dots, x_n | \theta)$$

Propiedades del estimador ML:

- Bajo condiciones de regularidad, es asintóticamente insesgado
- Es consistente
- Es asintóticamente eficiente (alcanza la cota de Cramér-Rao cuando $n \rightarrow \infty$)
- Es invariante bajo transformaciones: si $\hat{\theta}_{ML}$ es el ML de θ , entonces $g(\hat{\theta}_{ML})$ es el ML de $g(\theta)$

Método de solución: Criterio de la primera derivada:

$$\left. \frac{\partial \ln L(x_1, \dots, x_n | \theta)}{\partial \theta} \right|_{\theta = \hat{\theta}_{ML}} = 0$$

Verificar que la segunda derivada es negativa para confirmar que es un máximo.

7. Resumen de Verificaciones para un Estimador

Para analizar completamente un estimador $\hat{\theta}$, verificar:

1. **Inssegamiento:** Calcular $\mathbb{E}[\hat{\theta}]$ y verificar si es igual a θ
2. **Consistencia:** Verificar que $\lim_{n \rightarrow \infty} \text{Var}(\hat{\theta}) = 0$ (si es insesgado)
3. **Varianza:** Calcular $\text{Var}(\hat{\theta})$ explícitamente
4. **Eficiencia:** Calcular $\mathcal{I}(\theta)$ y comparar $\text{Var}(\hat{\theta})$ con $\frac{1}{\mathcal{I}(\theta)}$

1. Considere el problema de estimar $\theta \in \mathbb{R}$ dada una observación $X \in \mathbb{R}$, se sabe que la distribución condicional de Θ dado X está dotada de la siguiente función de densidad condicional definida como:

$$f_{\Theta|X}(\theta|x) = \begin{cases} e^{-(\theta-x)}, & \text{si } \theta > x \\ 0, & \text{si } x > \theta \end{cases}$$

Encuentre el estimador de mínimo error cuadrático medio y MAP.

Solución:

Para este problema, tenemos la distribución a posteriori $f_{\Theta|X}(\theta|x)$ y debemos encontrar dos estimadores distintos.

• **Estimador de Mínimo Error Cuadrático Medio (MMSE):**

El estimador MMSE corresponde a la esperanza condicional:

$$\hat{\theta}_{MMSE}(x) = \mathbb{E}[\Theta|X = x] = \int_{-\infty}^{\infty} \theta \cdot f_{\Theta|X}(\theta|x) d\theta$$

Dada la función de densidad:

$$f_{\Theta|X}(\theta|x) = \begin{cases} e^{-(\theta-x)}, & \text{si } \theta > x \\ 0, & \text{si } \theta \leq x \end{cases}$$

Entonces, la integral se reduce a:

$$\hat{\theta}_{MMSE}(x) = \int_x^{\infty} \theta \cdot e^{-(\theta-x)} d\theta$$

Realizamos el cambio de variable $u = \theta - x$, entonces:

- $\theta = u + x$
- $d\theta = du$
- Cuando $\theta = x$, tenemos $u = 0$
- Cuando $\theta \rightarrow \infty$, tenemos $u \rightarrow \infty$

Sustituyendo en la integral:

$$\hat{\theta}_{MMSE}(x) = \int_0^{\infty} (u + x) e^{-u} du$$

Separamos la integral en dos términos:

$$= \int_0^{\infty} u \cdot e^{-u} du + \int_0^{\infty} x \cdot e^{-u} du$$

Como x es una constante (no depende de u), podemos sacarlo de la integral:

$$= \int_0^{\infty} u \cdot e^{-u} du + x \cdot \int_0^{\infty} e^{-u} du$$

Ahora calculamos cada integral por separado:

Primera integral: $\int_0^\infty e^{-u} du$

Esta es una integral estándar de la exponencial:

$$\int_0^\infty e^{-u} du = \left[-e^{-u} \right]_0^\infty = \lim_{u \rightarrow \infty} (-e^{-u}) - (-e^0) = 0 - (-1) = 1$$

Segunda integral: $\int_0^\infty u \cdot e^{-u} du$

Usamos integración por partes. Recordemos que $\int v dw = vw - \int w dv$.

Elegimos:

- $v = u \implies dv = du$
- $dw = e^{-u} du \implies w = -e^{-u}$

Aplicando la fórmula:

$$\int_0^\infty u \cdot e^{-u} du = \left[u \cdot (-e^{-u}) \right]_0^\infty - \int_0^\infty (-e^{-u}) du$$

Simplificando:

$$= \left[-u \cdot e^{-u} \right]_0^\infty + \int_0^\infty e^{-u} du$$

Evaluamos el primer término:

$$\left[-u \cdot e^{-u} \right]_0^\infty = \lim_{u \rightarrow \infty} (-u \cdot e^{-u}) - (0 \cdot e^0) = 0 - 0 = 0$$

Nota: $\lim_{u \rightarrow \infty} u \cdot e^{-u} = 0$ porque la exponencial decrece más rápido que u crece (aplicando L'Hôpital si es necesario).

El segundo término ya lo calculamos antes y vale 1. Por lo tanto:

$$\int_0^\infty u \cdot e^{-u} du = 0 + 1 = 1$$

Combinando ambas integrales:

$$\hat{\theta}_{MMSE}(x) = \int_0^\infty u \cdot e^{-u} du + x \cdot \int_0^\infty e^{-u} du = 1 + x \cdot 1 = x + 1$$

Por lo tanto, el estimador MMSE es:

$\hat{\theta}_{MMSE}(x) = x + 1$

• **Estimador MAP (Maximum A Posteriori):**

El estimador MAP maximiza la densidad a posteriori:

$$\hat{\theta}_{MAP}(x) = \arg \max_{\theta} f_{\Theta|X}(\theta|x)$$

Recordemos que la función de densidad es:

$$f_{\Theta|X}(\theta|x) = \begin{cases} e^{-(\theta-x)}, & \text{si } \theta > x \\ 0, & \text{si } \theta \leq x \end{cases}$$

Para encontrar el máximo, analizamos el comportamiento de $f_{\Theta|X}(\theta|x) = e^{-(\theta-x)}$ en la región donde $\theta > x$.

Calculamos la derivada con respecto a θ :

$$\frac{d}{d\theta} f_{\Theta|X}(\theta|x) = \frac{d}{d\theta} e^{-(\theta-x)} = -e^{-(\theta-x)}$$

Como $e^{-(\theta-x)} > 0$ para todo θ , tenemos que:

$$\frac{d}{d\theta} f_{\Theta|X}(\theta|x) = -e^{-(\theta-x)} < 0 \quad \text{para todo } \theta > x$$

La derivada es siempre negativa en la región donde la función está definida, lo que significa que $f_{\Theta|X}(\theta|x)$ es una función estrictamente decreciente.

Por lo tanto, el máximo de $f_{\Theta|X}(\theta|x)$ se alcanza en el valor más pequeño posible de θ , que es $\theta = x$ (el límite inferior de la región donde la función es no nula).

Evaluando en ese punto:

$$f_{\Theta|X}(x|x) = e^{-(x-x)} = e^0 = 1$$

Y para cualquier $\theta > x$:

$$f_{\Theta|X}(\theta|x) = e^{-(\theta-x)} < 1$$

Concluimos que el estimador MAP es:

$$\hat{\theta}_{MAP}(x) = x$$

Interpretación:

- El estimador MAP $\hat{\theta}_{MAP}(x) = x$ elige el valor que maximiza la probabilidad a posteriori.
- El estimador MMSE $\hat{\theta}_{MMSE}(x) = x+1$ considera toda la distribución a posteriori y minimiza el error cuadrático medio esperado, resultando en un valor desplazado una unidad respecto al MAP debido a la forma asimétrica de la distribución exponencial.

2. Sean X_1, \dots, X_n muestras i.i.d. de la variable aleatoria $X \sim \mathcal{N}(\mu, \sigma^2)$. A su vez, μ es una variable aleatoria continua tal que $\mu \sim \mathcal{N}(0, \sigma_0^2)$.

1. Determine las densidades $f_{X_1, \dots, X_n}(x_1, \dots, x_n | \theta)$ y f_μ .
2. Obtenga el estimador MAP de μ , $\hat{\mu}_{MAP}$.
3. Determine los casos límite en que $\sigma_0^2 \rightarrow \infty$ y $\sigma_0^2 \rightarrow 0$. Interprete cada situación.

Solución:

1. Determine las densidades $f_{X_1, \dots, X_n}(x_1, \dots, x_n | \theta)$ y f_μ .

Como se ha visto en auxiliares anteriores, se tiene la siguiente densidad conjunta condicional para el vector X :

$$f_X(x_1, \dots, x_n | \mu) = \prod_{i=1}^n f_{X_i}(x_i | \mu) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}$$

$$f_X(x_1, \dots, x_n | \mu) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2}$$

Por otro lado, la densidad de μ es una normal conocida:

$$f_\mu(\mu) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \cdot e^{-\frac{\mu^2}{2\sigma_0^2}}$$

2. Obtenga el estimador MAP de μ , $\hat{\mu}_{MAP}$.

Por definición, se tiene que el estimador MAP es:

$$\hat{\mu}_{MAP} = \arg \max_{\mu} \ln[f_X(x_1, \dots, x_n | \mu) \cdot f_\mu(\mu)]$$

Obtengamos la expresión a maximizar:

$$\ln[f_X(x_1, \dots, x_n | \mu) \cdot f_\mu(\mu)] = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^n (x_i - \mu)^2 - \frac{1}{2} \ln(2\pi\sigma_0^2) - \frac{\mu^2}{2\sigma_0^2}$$

Luego, aplicamos el criterio de la primera derivada:

$$\begin{aligned}
 \frac{d}{d\mu} \left(-\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^n (x_i - \mu)^2 - \frac{1}{2} \ln(2\pi\sigma_0^2) - \frac{\mu^2}{2\sigma_0^2} \right) &= 0 \\
 \frac{1}{\sigma^2} \cdot \sum_{i=1}^n (x_i - \mu) - \frac{\mu}{\sigma_0^2} &= 0 \\
 \frac{1}{\sigma^2} \sum_{i=1}^n x_i - \frac{n\mu}{\sigma^2} - \frac{\mu}{\sigma_0^2} &= 0 \\
 \frac{\sum_{i=1}^n x_i}{\sigma^2} &= \mu \cdot \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2} \right) \\
 \mu &= \frac{\sum_{i=1}^n x_i}{\sigma^2 \cdot \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2} \right)} \\
 \mu &= \frac{\sum_{i=1}^n x_i}{\left(\frac{\sigma^2}{\sigma_0^2} + n \right)} \\
 \iff \mu &= \frac{\sigma_0^2}{n\sigma_0^2 + \sigma^2} \cdot \sum_{i=1}^n x_i \\
 \implies \mu &= \frac{\sigma_0^2}{\sigma_0^2 + \frac{\sigma^2}{n}} \cdot \frac{1}{n} \cdot \sum_{i=1}^n x_i
 \end{aligned}$$

Luego, el estimador MAP de μ corresponde a:

$$\hat{\mu}_{MAP} = \frac{\sigma_0^2}{\sigma_0^2 + \frac{\sigma^2}{n}} \cdot \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

3. Determine los casos límite en que $\sigma_0^2 \rightarrow \infty$ y $\sigma_0^2 \rightarrow 0$. Interprete cada situación.

Nos pondremos en ambos casos:

- $\sigma_0^2 \rightarrow \infty$: En este caso tenemos que ver el límite del estimador:

$$\lim_{\sigma_0^2 \rightarrow \infty} \frac{\sigma_0^2}{\sigma_0^2 + \frac{\sigma^2}{n}} \cdot \frac{1}{n} \cdot \sum_{i=1}^n x_i = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

Podemos notar de este caso que, mientras más incerteza tenemos sobre el parámetro μ en su distribución (equivalente a que la varianza de la normal sea muy grande), el estimador MAP tiende a la media empírica de las observaciones, es decir, los datos son los que determinan el estimador (Porque le creemos más a los datos que al conocimiento a priori). A su vez corresponde al estimador de máxima verosimilitud del problema.

- $\sigma_0^2 \rightarrow 0$: En este caso tenemos que ver el límite del estimador:

$$\lim_{\sigma_0^2 \rightarrow 0} \frac{\sigma_0^2}{\sigma_0^2 + \frac{\sigma^2}{n}} \cdot \frac{1}{n} \cdot \sum_{i=1}^n x_i = \frac{1}{\sigma^2} \cdot \sum_{i=1}^n x_i$$

En este caso ocurre que tenemos mucha certeza sobre el valor del estimador lo cual se modela como $\sigma_0^2 \rightarrow 0$. Esto genera que el estimador MAP quede dependa solo del número de observaciones y de la varianza de las observaciones. Esto se traduce a que se tiene con tanta certeza el valor del estimador MAP, que si $\sigma_0^2 \rightarrow 0$ entonces su valor siempre será su esperanza, en este caso 0. El estimador MAP sí entrega un valor aproximado pero como se tiene certeza total no es necesario.

3. Considere que tiene un cuerpo radiactivo el cual, dentro de un cierto intervalo de tiempo, emite Θ partículas. Sin embargo, usted mide cada una de estas emisiones con un detector imperfecto, el cual tiene probabilidad p de detectar correctamente cada partícula.

Si B_i^Θ son las variables aleatorias que indican si cada partícula es detectada correctamente o no, de modo que $B_i \sim \text{Bernoulli}(p)$, entonces la variable de medición X corresponde a

$$X = \sum_{i=1}^{\Theta} B_i,$$

correspondiente al total de partículas que fueron correctamente detectadas.

Gracias a estudios previos, usted sabe que $\Theta \sim \text{Poisson}(\lambda)$, con λ una propiedad del cuerpo radiactivo. Con esto, determine el estimador de mínimo error cuadrático medio (MMSE) para estimar el número de partículas a partir de las mediciones.

Solución:

Sabemos que el estimador MMSE $\hat{\theta}_{MMSE}(x)$ está dado por

$$\hat{\theta}_{MMSE}(x) = \mathbb{E}\{\Theta|X=x\},$$

por lo que para encontrarlo necesitamos expresar la distribución a posteriori. Para esto, notemos que por Bayes

$$P_{\Theta|X}(\theta|x) = \frac{P_{X|\Theta}(x|\theta)P_{\Theta}(\theta)}{P_X(x)},$$

donde

$$P_{X|\Theta}(x|\theta) = \binom{\theta}{x} p^x (1-p)^{\theta-x}, \quad P_{\Theta}(\theta) = \frac{e^{-\lambda} \lambda^\theta}{\theta!}.$$

Para el término del denominador, aplicamos la **ley de probabilidades totales**, que nos dice que para calcular la probabilidad marginal de X debemos sumar sobre todos los posibles valores de Θ :

$$P_X(x) = \sum_{\theta \in \mathbb{N}} P_{X|\Theta}(x|\theta)P_{\Theta}(\theta) = \sum_{\theta \geq 0} P_{X|\Theta}(x|\theta)P_{\Theta}(\theta)$$

Ahora bien, recordemos que $X|\Theta$ sigue una distribución binomial: $X|\Theta = \theta \sim \text{Binomial}(\theta, p)$, donde X representa el número de éxitos (partículas detectadas) en θ intentos. Por la naturaleza de la binomial, es imposible obtener más éxitos que intentos, es decir, $X \leq \Theta$. Esto significa que:

$$P_{X|\Theta}(x|\theta) = \binom{\theta}{x} p^x (1-p)^{\theta-x} = 0 \quad \text{si } \theta < x$$

porque el coeficiente binomial $\binom{\theta}{x} = \frac{\theta!}{x!(\theta-x)!}$ no está definido (o es cero) cuando $\theta < x$ (ya que $(\theta-x)!$ sería el factorial de un número negativo, que no existe). Por lo tanto, todos los términos con $\theta < x$ no contribuyen a la suma, y podemos escribir:

$$\begin{aligned} P_X(x) &= \sum_{\theta \geq 0} P_{X|\Theta}(x|\theta) P_{\Theta}(\theta) \\ &= \sum_{\theta \geq x} P_{X|\Theta}(x|\theta) P_{\Theta}(\theta) \\ &= \sum_{\theta \geq x} \binom{\theta}{x} p^x (1-p)^{\theta-x} \frac{e^{-\lambda} \lambda^{\theta}}{\theta!} \\ &= \sum_{\theta \geq x} \frac{\theta!}{x!(\theta-x)!} p^x (1-p)^{\theta-x} \frac{e^{-\lambda} \lambda^{\theta}}{\theta!} \\ &= \sum_{\theta \geq x} \frac{1}{x!(\theta-x)!} p^x (1-p)^{\theta-x} e^{-\lambda} \lambda^{\theta} \\ &= \frac{e^{-\lambda} p^x}{x!} \sum_{\theta \geq x} \frac{1}{(\theta-x)!} (1-p)^{\theta-x} \lambda^{\theta} \\ &= \frac{e^{-\lambda} (\lambda p)^x}{x!} \sum_{\theta \geq x} \frac{(\lambda(1-p))^{\theta-x}}{(\theta-x)!}, \end{aligned}$$

donde, considerando el cambio de variable $i = \theta - x$, podemos ver que la serie corresponde a $e^{\lambda(1-p)} = e^{\lambda} e^{-\lambda p}$, por lo que tenemos

$$P_X(x) = \frac{e^{-\lambda p} (\lambda p)^x}{x!},$$

indicando que $X \sim \text{Poisson}(\lambda p)$. Reemplazando, vemos que la distribución a posteriori (limitándose al rango $\theta \geq x$) corresponde a

$$\begin{aligned} P_{\Theta|X}(\theta|x) &= \frac{P_{X|\Theta}(x|\theta) P_{\Theta}(\theta)}{P_X(x)} \\ &= \frac{\frac{\theta!}{x!(\theta-x)!} p^x (1-p)^{\theta-x} \frac{e^{-\lambda} \lambda^{\theta}}{\theta!}}{\frac{e^{-\lambda p} (\lambda p)^x}{x!}} \\ &= \frac{e^{-\lambda(1-p)} (\lambda(1-p))^{\theta-x}}{(\theta-x)!}, \end{aligned}$$

correspondiente a una Poisson desplazada. Luego, calculando la esperanza tenemos

$$\begin{aligned}
 \hat{\theta}_{MMSE}(x) &= \mathbb{E}\{\Theta|X=x\} \\
 &= \sum_{\theta \geq x} \theta P_{\Theta|X}(\theta|x) \\
 &= \sum_{\theta \geq x} \theta \frac{e^{-\lambda(1-p)} (\lambda(1-p))^{\theta-x}}{(\theta-x)!} \quad / i = \theta - x \\
 &= e^{-\lambda(1-p)} \sum_{i \geq 0} (i+x) \frac{(\lambda(1-p))^i}{i!} \\
 &= e^{-\lambda(1-p)} \left[\sum_{i \geq 0} i \frac{(\lambda(1-p))^i}{i!} + x \sum_{i \geq 0} \frac{(\lambda(1-p))^i}{i!} \right].
 \end{aligned}$$

Cálculo del primer término: Para la suma $\sum_{i \geq 0} i \frac{(\lambda(1-p))^i}{i!}$, notamos que el término con $i = 0$ es cero, por lo que podemos escribir:

$$\begin{aligned}
 \sum_{i \geq 0} i \frac{(\lambda(1-p))^i}{i!} &= \sum_{i \geq 1} i \frac{(\lambda(1-p))^i}{i!} \\
 &= \sum_{i \geq 1} i \frac{(\lambda(1-p))^i}{i \cdot (i-1)!} \quad (\text{usando que } i! = i \cdot (i-1)!) \\
 &= \sum_{i \geq 1} \frac{(\lambda(1-p))^i}{(i-1)!}.
 \end{aligned}$$

Ahora, realizamos el cambio de variable $j = i - 1$, lo cual implica que $i = j + 1$. Cuando $i = 1$, tenemos $j = 0$, y cuando $i \rightarrow \infty$, tenemos $j \rightarrow \infty$. Además, $(\lambda(1-p))^i = (\lambda(1-p))^{j+1} = \lambda(1-p) \cdot (\lambda(1-p))^j$. Por lo tanto:

$$\begin{aligned}
 \sum_{i \geq 1} \frac{(\lambda(1-p))^i}{(i-1)!} &= \sum_{j \geq 0} \frac{(\lambda(1-p))^{j+1}}{j!} \\
 &= \lambda(1-p) \sum_{j \geq 0} \frac{(\lambda(1-p))^j}{j!} \\
 &= \lambda(1-p) \cdot e^{\lambda(1-p)}.
 \end{aligned}$$

Cálculo del segundo término: Para la suma $\sum_{i \geq 0} \frac{(\lambda(1-p))^i}{i!}$, reconocemos la serie de Taylor de la exponencial:

$$\sum_{i \geq 0} \frac{(\lambda(1-p))^i}{i!} = e^{\lambda(1-p)}.$$

Combinando ambos términos: Reemplazando en la expresión original:

$$\begin{aligned}\hat{\theta}_{MMSE}(x) &= e^{-\lambda(1-p)} \left[\lambda(1-p) \cdot e^{\lambda(1-p)} + x \cdot e^{\lambda(1-p)} \right] \\ &= e^{-\lambda(1-p)} \cdot e^{\lambda(1-p)} [\lambda(1-p) + x] \\ &= \lambda(1-p) + x \\ &= x + \lambda(1-p),\end{aligned}$$

de modo que el estimador MMSE está dado por

$$\boxed{\hat{\theta}_{MMSE}(x) = x + \lambda(1-p).}$$

Podemos ver que, pese a que $\Theta \in \mathcal{A} = \mathbb{N}$ (es decir, el problema es de detección), dado que $\lambda \in \mathbb{R}^+$ y $p \in (0, 1)$ el estimador elige valores dentro de \mathbb{R}^+ , y por lo tanto el problema se trata, indirectamente, como uno de estimación.