Resolución Segundo Parcial

Métodos Estadísticos en Física Experimental

16 Marzo 2019

Problema 1

Sea la función distribución Gamma$(x|\alpha, \beta)$.

a) Considerando α conocido, encontrar el estimador de máxima verosimilitud para β.

b) ¿Es 100% eficiente?, y si no lo es, buscar una función de él que si lo sea.

c) Calcular el sesgo y la varianza del estimador 100% eficiente hallado en el ítem anterior.

Resolución 1

a) $\mathcal{L}(x;\theta) = \prod_{i=1}^{n} \frac{\beta^{\alpha}}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\beta x_i} \implies ln\mathcal{L} = \sum_{i=1}^{n}[\alpha ln\beta + (\alpha-1)lnx_i - \beta x_i - ln\Gamma(\alpha)]$

$\frac{\partial ln\mathcal{L}}{\partial\beta} = \sum_{i=1}^{n}(\frac{\alpha}{\beta} - x_i) = \frac{n\alpha}{\beta} - \sum_{i=1}^{n} x_i = 0 \implies \hat{\beta} = \frac{n\alpha}{\sum_{i=1}^{n} x_i}$

b) Quiero ver si puedo escribir $\frac{\partial ln\mathcal{L}}{\partial\beta} = A(\alpha,\beta)[t(\vec{x}) - h(\beta) - b(\beta)]$.

Efectivamente, $\frac{\partial ln\mathcal{L}}{\partial\beta} = \frac{n\alpha}{\beta} - \sum_{i=1}^{n} x_i = -n\alpha\left(\frac{\sum_{i=1}^{n} x_i}{n\alpha} - \frac{1}{\beta}\right)$.

Entonces, $t = \frac{\sum_{i=1}^{n} x_i}{n\alpha}$ es un estimador 100% eficiente de la función $\frac{1}{\beta}$.

c) Del item anterior, salió directamente que el sesgo $b(\beta) = 0$.

Para calcular la varianza del estimador 100% eficiente usamos el límite inferior de Cramer-Rao:

$Var(t) = \left|\frac{\frac{\partial h}{\partial\beta} + \frac{\partial b}{\partial\beta}}{A(\alpha,\beta)}\right| = \left|\frac{-\frac{1}{\beta^2}}{-n\alpha}\right| = \frac{1}{n\alpha\beta^2}$

Problema 2

Considerar el experimento que consiste en contar el número de trazas de muones que se observan en un minuto utilizando una pequeña cámara de niebla.

a) ¿Cuál fue el número de muones observados si el intervalo frecuentista reportado para el flujo de muones por minuto con un 90% de nivel de confianza es (0, 7.99)?

b) ¿Cual habría sido el intervalo informado si se hubiese decidido por uno central frecuentista con el 68% de nivel de confianza?

Resolución 2

a) Dado el intervalo (0, 7.99) sabemos que se trata de un intervalo con cota superior, y que dicha cota es μ_U=7.99. La pregunta que debemos responder para saber cual fue el número de muones observados es: desde 0 hasta qué valor hubo que sumar términos de la distribucion $P(k|\mu_U)$ para que de 0.10? Esto es, pedir que en la cola izquierda de la distribución del estadístico $t = k$ quede el 10%. Mirando la última columna de la tabla encuentro que $\Sigma_{k=0}^{k=4} P(k|\mu_U)$=0.100, por lo tanto, podemos inferir que se observaron $k = 4$ trazas de muones.

b) Para informar el intervalo central frecuentista con un 68% de nivel de confianza, es necesario dejar un 16% a izquierda y a derecha. Entonces, sabiendo ahora que la suma debe ser hasta $k = 4$, busco en la tabla, para que valor de μ_U se cumple que $\Sigma_{k=0}^{k=4} P(k|\mu_U)$=0.16. Para encontrar μ_L busco que $\Sigma_{k=4}^{k=\infty} P(k|\mu_L)$=0.16, lo que equivale a pedir que, $\Sigma_{k=0}^{k=3} P(k|\mu_L)$=0.84. Esto se cumple para $\mu_U = 7.14$ y $\mu_L = 2.09$.

Problema 3

Para el mismo experimento del problema anterior, ahora se quiere informar un intervalo utilizando el enfoque bayesiano.

a) Encontrar la relación de transformación entre el Prior y el Posterior sabiendo que la distribución Gamma$(\theta|\alpha, \beta)$ es la conjugada de la distribución de Poisson$(k|\mu)$.

b) Se realiza una medición de un minuto de duración y se observan cinco trazas de muones. Usando la distribución Gamma$(\theta|\alpha = 7, \beta = 2.2)$ como prior, indique el estimador bayesiano para μ.

c) Informar un intervalo para μ usando como receta para su construcción: E$(x) \pm \sqrt{Var(x)}$.

d) Calcular la cobertura de la receta del ítem anterior para $\mu = 4$. ¿Se parece al 68% que hubiese esperado antes de cursar esta materia?

Resolución 3

a) Sabemos que la likelihood es Poissoniana: $\mathcal{L}(\vec{k}, \mu) = \prod_{i=1}^{n} \frac{e^{-\mu}\mu^{k_i}}{k!}$. Llamando $\bar{k} = \frac{\sum_{i=1}^{n} k_i}{n}$, podemos reescribirla como: $\mathcal{L}(\vec{k}, \mu) = \frac{e^{-n\mu}\mu^{n\bar{k}}}{\prod_{i=1}^{n} k_i!}$.

Nos proponen como prior: $\Pi(\mu) = Gamma(\mu; \alpha, \beta) = \frac{\beta^\alpha \mu^{\alpha-1} e^{-\beta\mu}}{\Gamma(\alpha)}$.

Con esto, ya podemos calcular el posterior como $f(\mu; \vec{k}) = \frac{\mathcal{L}(\vec{k},\mu)\Pi(\mu)}{\int \mathcal{L}(\vec{k},\mu)\Pi(\mu)d\mu}$.

$$f(\mu, \vec{k}) = \frac{\frac{e^{-n\mu}\mu^{n\bar{k}}}{\prod_{i=1}^{n} k_i!} \frac{\beta^\alpha \mu^{\alpha-1} e^{-\beta\mu}}{\Gamma(\alpha)}}{\int_0^\infty \frac{e^{-n\mu}\mu^{n\bar{k}}}{\prod_{i=1}^{n} k_i!} \frac{\beta^\alpha \mu^{\alpha-1} e^{-\beta\mu}}{\Gamma(\alpha)} d\mu} = \frac{e^{-\mu(n+\beta)}\mu^{n\bar{k}+\alpha-1}}{\int_0^\infty e^{-\mu(n+\beta)}\mu^{n\bar{k}+\alpha-1} d\mu}.$$

Para que la integral del denominador quede correctamente normalizada, hay que multiplicar y dividir por el factor $\frac{(n+\beta)^{n\bar{k}+\alpha}}{\Gamma(n\bar{k}+\alpha)}$. Así, la integral da 1 y el posterior resulta:

$$f(\mu, \vec{k}) = (n + \beta)^{n\bar{k}+\alpha} \frac{e^{-\mu(n+\beta)}\mu^{n\bar{k}+\alpha-1}}{\Gamma(n\bar{k}+\alpha)}$$

b) Ahora tenemos $\alpha = 7$, $\beta = 2.2$, $n = 1$, $k = 5$. Con estos datos, podemos calcular la esperanza del posterior y obtenemos $\hat{\mu} = 3.75$.

c) Para dar el intervalo para μ, sólo nos falta calcuar la varianza del posterior, que se hace con los mismos datos del item anterior y arroja un resultado de $Var(\hat{\mu}) = 1.16$. Con esto, el intervalo resulta: $[2.67, 4.83]$.

d) Para calcular la cobertura de esta receta necesito saber que fracción de los intervalos que ella genera incluirán al valor real. Calculo entonces los intervalos que la receta devuelve para diferentes realizaciones de la variable aleatoria k. Asi es que: $k = 0[1,4;3,0]$, $k = 1[1,6;3,4]$, $k = 2[1,9;3,8]$, $k = 3[2,1;4,1]$, $k = 4[2,4;4,5]$, $k = 5[2,7;4,8]$, $k = 6[2,9;5,2]$, $k = 7[3,2;5,5]$, $k = 8[3,5;5,9]$, $k = 9[3,8;6,3]$, $k = 10[4,0;6,6]$ Y de todos ellos, solo los intervalos obtenidos para valores de k entre 3 y 9 incluyen a $\mu = 4$. Para dar la cobertura solo me resta sumar la probabilidad de ocurrencia de cada uno de los intervalos que si incluyen a μ y para eso puedo usar las probabilidades de la cuarta columna de la tabla. Asi obtengo que la cobertura es del 75,4%. Muy lejos del 68% que antes de cursar MEFE hubieramos sospechado.

Quienes hayan usado los parámetros del posterior hallados en el item b), obtuvieron otros intervalos y una cobertura menor, pero en tanto el planeto haya sido este, la resolución se considerará correcta.

Problema 4

Con un experimento similar al del problema 2, pero de mayor duración por medición, se quiere establecer una alarma que avise cuando el promedio del conteo de muones resulta sospechosamente alto, siendo el valor de μ_0 esperado igual a 100.

a) Determinar el valor de k_{max} de la alarma para establecer un test de significancia α=0.025.

b) Calcular la potencia del test considerando μ_1=169 como hipótesis alternativa.

Resolución 4

a) Considerando que μ_0 es grande, podemos aproximar la distribución como una gaussiana $N(\mu_0, \sigma_0 = \sqrt{\mu_0}$ y resulta que $\alpha = 0.025$ es la cola a la derecha de $\mu_0 + 2\sigma_0$. Con lo cual, el $k_{max} = \mu_0 + 2\sigma_0 = 120$.

b) La potencia se calcula como $\int_{k_{max}}^{\infty} N(\mu_1, \sigma_1 = \sqrt{\mu_1})dx$. Pero como no sabemos hacer esta cuenta vamos a ponerle una cota. Como $\mu_1 - 3\sigma_1 > k_{max}$, podemos decir que la potencia del test seguro es mayor que 99%.

Datos útiles: Gamma$(x|\alpha,\beta)$=$\beta^\alpha x^{\alpha-1}\exp(-\beta x)/\Gamma(\alpha)$, E$(x)$=$\alpha/\beta$ y Var(x)=α/β^2.

	μ_1=2.09		μ_2=4.00		μ_3=7.14		μ_4=7.99	
k^*	$P(k^*\|\mu_1)$	$\Sigma_{k=0}^{k*}P(k\|\mu_1)$	$P(k^*\|\mu_2)$	$\Sigma_{k=0}^{k*}P(k\|\mu_2)$	$P(k^*\|\mu_3)$	$\Sigma_{k=0}^{k*}P(k\|\mu_3)$	$P(k^*\|\mu_4)$	$\Sigma_{k=0}^{k*}P(k\|\mu_4)$
0	0.12369	0.12369	0.01832	0.01832	0.00079	0.00079	0.00034	0.00034
1	0.25851	0.38219	0.07326	0.09158	0.00566	0.00645	0.00271	0.00305
2	0.27014	0.65233	0.14653	0.23810	0.02021	0.02666	0.01082	0.01386
3	0.18820	0.84053	0.19537	0.43347	0.04809	0.07475	0.02881	0.04267
4	0.09833	0.93886	0.19537	0.62884	0.08585	0.16060	0.05754	0.10021
5	0.04110	0.97996	0.15629	0.78513	0.12259	0.28319	0.09195	0.19215
6	0.01432	0.99428	0.10420	0.88933	0.14588	0.42907	0.12244	0.31460
7	0.00427	0.99856	0.05954	0.94887	0.14880	0.57786	0.13976	0.45436
8	0.00112	0.99967	0.02977	0.97864	0.13280	0.71066	0.13959	0.59394
9	0.00026	0.99993	0.01323	0.99187	0.10536	0.81602	0.12392	0.71786
10	0.00005	0.99999	0.00529	0.99716	0.07522	0.89124	0.09901	0.81688
11	0.00001	1.00000	0.00192	0.99908	0.04883	0.94007	0.07192	0.88880
12	0.00000	1.00000	0.00064	0.99973	0.02905	0.96912	0.04789	0.93668