Murgio, Maximiliano
DNI: 42496527

2do Parcial MEFE
LU: 36/20

10 (diez)
Impecable!

1/5
23/06

- 1er BLOQUE:

1) Gatubela le propone a Batman que haga el experimento de una binomial negativa con $k=1$, ya que debe contar murciélagos hasta que dé con uno que tenga el Batianillo. Como lo debe hacer por 3 noches, esto requiere de realizar el experimento 3 veces. La verosimilitud se escribe como:

$$\mathcal{L}(\underline{x}|p) = \prod_{i=1}^{3} BN(x_i|k_i,p)$$

$$\mathcal{L}(\underline{x}|p) = \prod_{i=1}^{3} \binom{x_i-1}{k_i-1} p^{k_i}(1-p)^{x_i-k_i}$$

pero, siempre $k=1$, por lo que $k_i=1\ \forall i$, y el combinatorio $\binom{x_i-1}{0}=1$, por lo que me queda:

$$\mathcal{L}(\underline{x}|p) = \prod_{i=1}^{3} p(1-p)^{x_i-1}$$

donde 'x_i' es el número de murciélagos que contó la i-ésima noche hasta encontrar uno con el Batianillo y 'p' la probabilidad de que un murciélago al azar en la ciudad lleve puesto el Batianillo

2) Si quiero un estadístico suficiente para 'p', debo ver si puedo escribir a mi $\mathcal{L}(\underline{x}|p)$ en la familia exponencial:

$$\mathcal{L}(\underline{x}|p) = \exp\left(\sum_i B_i(p)C_i(\underline{x}) + D(p) + Z(\underline{x})\right)$$

si es así, entonces los $t_i(\underline{x}) = \sum_i C_i(\underline{x})$ son estadísticos suficien-

TES PARA p (CONDICIÓN DE DARMOIS). PONGAMOS ESTO EN PRÁCTICA:

$$\mathcal{L}(\underline{x}|p) = \prod_{i=1}^{3} p(1-p)^{x_i-1}$$

$$= \exp\left(\ln\left(\prod_{i=1}^{3} p(1-p)^{x_i-1}\right)\right)$$

$$= \exp\left(\sum_{i=1}^{3} \ln\left(p(1-p)^{x_i-1}\right)\right)$$

$$= \exp\left(\sum_{i=1}^{3} \ln(p) + \sum_{i=1}^{3}(x_i-1)\ln(1-p)\right)$$

$$= \exp\left(\underbrace{\ln(1-p)}_{B(p)}\underbrace{\sum_{i=1}^{3} x_i}_{C(\underline{x})} + \underbrace{3\ln(p) - 3}_{D(p)}\right)$$

POR LO QUE PODEMOS VER QUE EXISTE UN ESTADÍSTICO SUFICIENTE $t(\underline{x})$ PARA LA PROBABILIDAD p, EL CUAL ES:

$$t(\underline{x}) = \sum_{i=1}^{3} x_i \rightarrow \text{ESTADÍSTICO SUFICIENTE.}$$

Y LO ES TAMBIÉN CUALQUIER FUNCIÓN $t' = f(t)$ BIYECTIVA.

3) VEAMOS SI EXISTE UN ESTIMADOR 100% EFICIENTE PARA 'p'. PARA ESO, TENEMOS QUE VER SI SE CUMPLE LA CONDICIÓN DE CRAMER RAO:

$$\frac{\partial \ln \mathcal{L}(\underline{x}|p)}{\partial p} = A(p)\left[T(\underline{x}) - h(p)\right]$$

SI ES ASÍ, EL ESTADÍSTICO $T(\underline{x})$ ES 100% EFICIENTE PARA '$h(p)$' Y TIENE VARIANZA MÍNIMA. DERIVEMOS Y VEAMOS QUÉ SUCEDE:

$$\mathcal{L}(\underline{x}|p) = \prod_{i=1}^{3} p(1-p)^{x_i-1}$$

$$\ln\left(\mathcal{L}(\underline{x}|p)\right) = \sum_{i=1}^{3} \ln(p) + \sum_{i=1}^{3}(x_i-1)\ln(1-p)$$

$$\frac{\partial \ln \mathcal{L}(\underline{x}|p)}{\partial p} = \frac{3}{p} - \frac{1}{1-p}\sum_{i=1}^{3}(x_i-1)$$

$$= \frac{3}{p} + \frac{3}{1-p} - \frac{1}{1-p}\sum_{i=1}^{3}x_i$$

$$= \frac{3(1-p)+3p}{p(1-p)} - \frac{1}{1-p}\cdot\sum_{i=1}^{3}x_i$$

$$= \frac{3}{p(1-p)} - \frac{1}{1-p}\sum_{i=1}^{3}x_i$$

$$= \underbrace{-\frac{3}{1-p}}_{A(p)}\left[\underbrace{\frac{\sum_{i=1}^{3}x_i}{3}}_{T(\bar{x})} - \underbrace{\frac{1}{p}}_{h(p)}\right] \rightarrow$$ CONDICIÓN DE CRAMER RAO.

ENTONCES, PODEMOS DECIR QUE NO EXISTE UN ESTIMADOR 100% EFICIENTE PARA p, PERO SÍ EXISTE PARA $h(p) = p^{-1}$, EL CUAL ES:

$$T(\underline{x}) = \frac{1}{3}\sum_{i=1}^{3}x_i = \frac{\sum_{i=1}^{3}x_i}{\sum_{i=1}^{3}k_i}$$

CALCULEMOS SU VARIANZA, LA CUAL SERÁ MÍNIMA POR SER 100% EFICIENTE (IGUALDAD EN LA COTA CRAMER RAO):

$$\mathrm{Var}(T(\underline{x})) = \frac{(\partial h/\partial p)^2}{E\left[-\frac{\partial^2 \ln \mathcal{L}}{\partial p^2}\right]}$$ → VARIANZA DE $T(\underline{x})$.

calculemos los términos involucrados:

- $h(p) = \frac{1}{p} \Rightarrow \frac{\partial h}{\partial p} = -\frac{1}{p^2}$

- $\frac{\partial^2 \ln \mathcal{L}}{\partial p^2} = -\frac{3}{p^2} - \frac{1}{(1-p)^2} \sum_{i=1}^{3} (x_i - 1)$

- $E\left[-\frac{\partial^2 \ln \mathcal{L}}{\partial p^2}\right] = E\left[\frac{3}{p^2} + \frac{1}{(1-p)^2} \sum_{i=1}^{3} (x_i - 1)\right]$

$$= \frac{3}{p^2} + \frac{1}{(1-p)^2} \sum_{i=1}^{3} E(x_i - 1)$$

$E(x_i) = \frac{k_i}{p} = \frac{1}{p}$

$$= \frac{3}{p^2} + \frac{1}{(1-p)^2} \sum_{i=1}^{3} \left(\frac{1}{p} - 1\right)$$

$$= \frac{3}{p^2} + \frac{3}{(1-p)^2} \frac{\cancel{(1-p)}}{p} = \frac{3}{p^2} + \frac{3}{p(1-p)}$$

ENTONCES:

$$\mathrm{Var}(T(\underline{x})) = \frac{1/p^4}{\frac{3}{p^2} + \frac{3}{p(1-p)}} = \frac{\cancel{p}^{1}(1-p)}{\cancel{p^4}^{2}\left[3(1-p) + 3p\right]}$$

$$\boxed{\mathrm{Var}(T(\underline{x})) = \frac{1-p}{3p^2}}$$ → VARIANZA DE MI ESTADÍSTICO 100% EFICIENTE $T(\underline{x})$.

Todo el Bloque 1 Bien y Muy Prolijo, No Voy a Negar cierta preocupación por la presión que Pones sobre el Lápiz // lapicera

• 2DO BLOQUE:

4) PARA OBTENER UNA ESTIMACIÓN FRECUENTISTA PARA LA PROBABILIDAD 'p' DE QUE UN MURCIÉLAGO AL AZAR EN LA CIUDAD LLEVE PUESTO EL BASTIANILLO, PODEMOS BUSCAR EL ESTIMADOR DE MÁXIMA VEROSIMILITUD $\hat{p}$.

REESCRIBO LA VEROSIMILITUD OBTENIDA EN EL BLOQUE ANTERIOR (n NOCHES):

$$\mathcal{L}(\underline{x}|p) = \prod_{i=1}^{n} p(1-p)^{x_i-1}$$, CON $n=1$ PORQUE HIZO UNA NOCHE.

SI APLICO $\ln()$, NO MODIFICO EL VALOR DE $\hat{p}$ AL QUERER BUSCAR EL MÁXIMO DE $\mathcal{L}$, POR SER $\ln()$ UNA FUNCIÓN CRECIENTE:

$$\ln(\mathcal{L}(\underline{x}|p)) = \ln(p) + (x_i-1)\ln(1-p)$$

AHORA, DERIVEMOS E IGUALAMOS A CERO:

$$\left.\frac{\partial \ln(\mathcal{L})}{\partial p}\right|_{\hat{p}} = 0 \Rightarrow \frac{1}{\hat{p}} - \frac{1}{1-\hat{p}}(x_i-1) = 0$$

$$1-\hat{p} = \hat{p}(x_i-1)$$

$$1 = \hat{p}(x_i-1) + \hat{p}$$

$$1 = \hat{p}.x_i \Rightarrow \boxed{\hat{p} = \frac{1}{x_i}}$$

DONDE x_i ES EL NÚMERO DE MURCIÉLAGOS QUE CONTÓ

NOS DICEN QUE $x_i = 70$ $\Rightarrow$ $\hat{p}_{FREC.} = \frac{1}{70}$ $\rightarrow$ ESTIMACIÓN FRECUENTISTA PARA p.

5) Ahora, agarremos el problema desde el enfoque Bayesiano. Me piden que use un prior uniforme, y como $p \in [0,1]$ debo usar la distribución $U_{[0,1]}$.

Me parece un prior razonable sabiendo que mi verosimilitud viene de un binomial negativa y la uniforme es un caso particular de la distribución Beta ($B(x|\alpha=1, \beta=1)$), por lo que estaríamos usando su distribución conjugada (siendo el posterior de la misma familia). Además, tenemos absoluto desconocimiento de la cantidad de murciélagos que hay en la ciudad a comparación de los que hay en la Baticueva, si bien uno podría intuir que la probabilidad será baja ($p << 1$) debido a la dimensionalidad de ambos, por lo que también podría usarse un prior que "beneficie" a p bajos. ¿se te ocurre alguno?

Dejando esto de lado, hagamos cuentas:

$$P(p) = \frac{\mathcal{L}(\underline{x}|p) \cdot \Pi(p)}{\int \mathcal{L}(\underline{x}|p)\Pi(p)\, dp} \underset{\Pi(p) = U_{[0,1]}}{\Longrightarrow} \boxed{P(p|\underline{x}) = \frac{\mathcal{L}(\underline{x}|p)}{\int_0^1 \mathcal{L}(\underline{x}|p)\, dp}}$$

$$\int_0^1 \mathcal{L}(\underline{x}|p)\, dp = \int_0^1 p \cdot (1-p)^{70-1}\, dp = \int_0^1 p(1-p)^{69}\, dp$$

$$= -\int_1^0 (1-u)\,u^{69}\, du = \int_0^1 (u^{69} - u^{70})\, du = \left[\frac{u^{70}}{70} - \frac{u^{71}}{71}\right]\Bigg|_0^1$$

$$= \frac{1}{70} - \frac{1}{71} = \frac{1}{4970}$$

Cambio de variable

$u = 1-p \rightarrow p = 1-u$

$du = -dp$ $\quad p \to 0 \Rightarrow u \to 1$

$\quad p \to 1 \Rightarrow u \to 0$

MURGIA,
MAXIMILIANO
DNI: 42490527

por lo tanto, el posterior queda como:

$$p(p|\underline{x}) = 4970\, p\, (1-p)^{69}$$

Otra forma hubiese sido tratar de ver la proporcionalidad del posterior con la $\mathcal{L}$ y el prior:

$$p(p|\underline{x}) \propto \mathcal{L}(\underline{x}|p)\, \Pi(p)$$

y como $\Pi(p) = U[0,1]$, sólo haría falta ver el factor de normalización de $\mathcal{L}$ al ver que se trata de la Beta $B(x|\alpha=2, \beta=70)$:

$$\mathcal{L}(\underline{x}|p) \propto B(\underline{x}|\alpha=2, \beta=70)$$

y se normaliza con:

$$p(p|\underline{x}) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot p^{\alpha-1}(1-p)^{\beta-1}$$

$$= \frac{\Gamma(2+70)}{\Gamma(2)\Gamma(70)} \cdot p\,(1-p)^{69} \quad , \quad \begin{array}{l} \Gamma(72) = 71! \\ \Gamma(2) = 1! = 1 \\ \Gamma(70) = 69! \end{array}$$

$$\Rightarrow p(p|\underline{x}) = \frac{71!}{69!}\, p(1-p)^{69}$$

$$p(p|\underline{x}) = 4970\, p\,(1-p)^{69} \rightarrow \text{MISMO RESULTADO.}$$

↳ BETA$(x|\alpha=2, \beta=70)$

Bien!

SI QUIERO UNA ESTIMACIÓN $\hat{p}$ BAYESIANO, PUEDO USAR LA ESPERANZA DE LA DISTRIBUCIÓN BETA $(\underline{X} \mid \alpha=2, \beta=70)$:

$$E(B(\alpha,\beta)) = \frac{\alpha}{\alpha+\beta} \Rightarrow E(B(2,70)) = \frac{2}{2+70}$$

$$\hat{p}_{Bay} = E(B(2,70)) = \frac{1}{36}$$ → ESTIMACIÓN BAYESIANA PARA p.

6) HACEMOS LA ESTIMACIÓN DEL NÚMERO TOTAL DE MURCIÉLAGOS EN CIUDAD GÓTICA, UTILIZANDO LOS $\hat{p}$ OBTENIDOS CON LOS ENFOQUES FRECUENTISTA Y BAYESIANO. AL SER DISTRIBUCIÓN BINOMIAL NEGATIVA:

- FRECUENTISTA: $K=290$, $\hat{p} = 1/70$.

$\hat{n} = K/\hat{p} \Rightarrow \hat{n}_{FREC} = 20.300$ murciélagos → ESTIMACIÓN CON ENFOQUE FRECUENTISTA

- BAYESIANO: $K=290$, $\hat{p} = 1/36$

$\hat{n} = K/\hat{p} \Rightarrow \hat{n}_{BAY} = 10.440$ murciélagos → ESTIMACIÓN CON ENFOQUE BAYESIANO

NOTA

MURGIA, MAXIMILIANO
DNI: 42498527

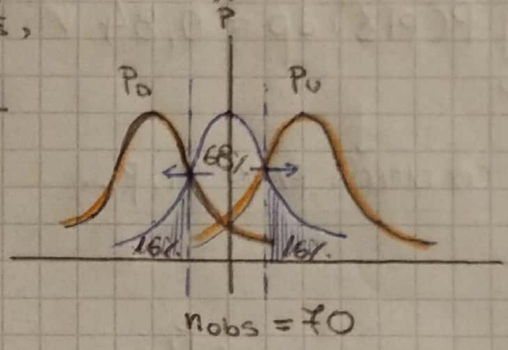
2$^{\text{DO}}$ PARCIAL MEFE
LU: 36/20

5/5
23/06

• 3$^{\text{ER}}$ BLOQUE:

7) SABEMOS QUE BATMAN TUVO 70 OBSERVACIONES REALIZADAS HASTA QUE PUDO ENCONTRAR 1 MURCIÉLAGO CON BATIANILLO. LO QUE AHORA DEBERÍAMOS HACER PARA HALLAR EL INTERVALO FRECUENTISTA CENTRAL DEL 68% ES OBSERVAR LAS CURVAS QUE NOS DIO ALFRED.

CUALITATIVAMENTE, CON LO QUE DIBUJÉ, DEBO VER EN LAS CURVAS PARA QUÉ VALOR DE p CONSIGO QUE LO OBSERVADO HASTA $n=69$ QUEDE EN LA COLA IZQUIERDA (APROX CON 16% DE CL) Y HASTA $n=70$ QUEDE A LA IZQUIERDA DE LA COLA DERECHA (APROX CON 16% DE CL). MATEMÁTICAMENTE ES:



• $$\sum_{n=0}^{69} p_U(1-p_U)^{n-1} = 0,16$$

• $$\sum_{n=71}^{+\infty} p_D(1-p_D)^{n-1} = 0,16 \Rightarrow \sum_{n=0}^{70} p_D(1-p_D)^{n-1} = 0,84$$

CONDICIONES PARA p_U Y p_D

SIENDO p_U EL LÍMITE SUPERIOR DE MI INTERVALO Y p_D EL LÍMITE INFERIOR DEL MISMO. MIRANDO LA TABLA, ME ENCUENTRO QUE EL INTERVALO FRECUENTISTA ES:

$$I_F = [0,0025 \,;\, 0,0260]$$ ✓

68% CL CENTRAL.

INTERVALO FRECUENTISTA

NOTA:

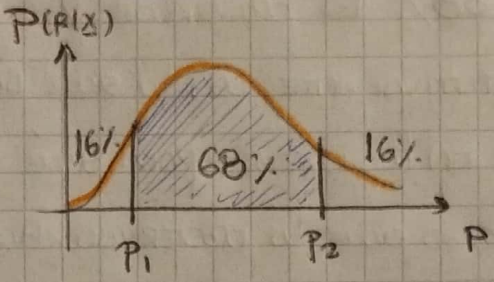
8) Para el intervalo Bayesiano, debo simplemente integrar en el posterior $P(p|\underline{x})$ que hallé antes y buscar en qué intervalo central con CL 68% se cumple. Por lo cual:

- $\displaystyle\int_0^{p_1} P(p|\underline{x})\,dp = 0{,}16$ ✓

- $\displaystyle\int_0^{p_2} P(p|\underline{x})\,dp = 0{,}84$ ✓

↓ condición para p_1, p_2.

, con $P(p,\underline{x}) = 4970\,p(1-p)^{69}$



$$\int_0^{p_1} P(p,x)\,dp = \int_0^{p_1} 4970\,p(1-p)^{69}\,dp = \int_0^{1-u_1} 4970\,(1-u)\,u^{69}\,du$$

$$= 4970 \int_0^{1-u_1} (u^{69} - u^{70})\,du = 4970\left(\frac{u^{70}}{70} - \frac{u^{71}}{71}\right)\Bigg|_0^{1-u_1} = 0{,}16$$

Cambio de variable

$1-p = u$, $u_1 = 1-p_1 \rightarrow p_1 = 1-u_1$

$dp = -du$

$$4970\left[\frac{(1-u_1)^{70}}{70} - \frac{(1-u_1)^{71}}{71}\right] = 0{,}16$$

$$4970\left[\frac{p_1^{70}}{70} - \frac{p_1^{71}}{71}\right] = 0{,}16$$

↓ Ecuación para p_1

y, para p_2, es análoga:

$$4970\left[\frac{p_2^{70}}{70} - \frac{p_2^{71}}{71}\right] = 0{,}84$$

↓ Ecuación para p_2

y el intervalo Bayesiano queda:

$$I_B = [p_1, p_2] \quad \text{68\% CL central}$$

↳ Intervalo Bayesiano.

Escaneado con CamScanner