SOLUCIONES EXAMEN DE RECUPERACIÓN (PROBABILIDAD Y ESTADÍSTICA)

a/ Una variable aleatoria continua R tiene la siguiente función densidad de probabilidad:

$$p(R) = \begin{cases} \frac{1}{2L}\,, & \text{si } 0 \leq R \leq L, \\ \frac{R}{2L\sqrt{R^2-L^2}}\,, & \text{si } L < R \leq L\sqrt{2}, \end{cases}$$

siendo L una constante positiva.

1) ¿Existe alguna restricción adicional sobre la constante L?

2) Hallar el valor medio de la variable aleatoria R.

1) Comprobamos si $p(R)$ cumple la condición de normalización

$$\int_0^{L\sqrt{2}} p(R)\,dR = \int_0^{L} \frac{1}{2L}\,dR + \frac{1}{4L}\int_L^{L\sqrt{2}} \frac{2R\,dR}{\sqrt{R^2-L^2}} = \frac{1}{2} + \frac{1}{4L}\left[2\left(R^2-L^2\right)^{\frac{1}{2}}\right]_L^{L\sqrt{2}} =$$

$$= \frac{1}{2} + \frac{1}{4L}\left[2L\right] = \frac{1}{2} + \frac{1}{2} = 1$$

La condición de normalización se cumple para todo valor de $L>0$ así que no existe ninguna restricción adicional sobre L

2) Aplicamos la definición de valor medio de una variable aleatoria

$$\langle R\rangle = \int_0^{L\sqrt{2}} R\,p(R)\,dR = \frac{1}{2L}\int_0^{L} R\,dR + \frac{1}{2L}\int_L^{L\sqrt{2}} \frac{R^2\,dR}{\left(R^2-L^2\right)^{\frac{1}{2}}} =$$

$$= \frac{1}{2L}\,\frac{L^2}{2} + \frac{L}{4}\left(ch^{-1}[\sqrt{2}] + \sqrt{2}\right) = \frac{L}{4}\left[1+\sqrt{2}+ch^{-1}[\sqrt{2}]\right] =$$

$$= \boxed{\frac{L}{4}\left(1+\sqrt{2}+\log[1+\sqrt{2}]\right)}$$

Usar que $ch^{-1}[\sqrt{2}] = \log[1+\sqrt{2}]$

Cálculo de la integral:

$$\frac{1}{2L}\int_{L}^{L\sqrt{2}} \frac{R^2\,dR}{(R^2-L^2)^{\frac{1}{2}}} = \begin{bmatrix} R = L\,ch\,x \\ dR = L\,sh\,x\,dx \end{bmatrix} = \frac{1}{2L}\int_{ch^{-1}[1]}^{ch^{-1}[\sqrt{2}]} \frac{L^2\,ch^2 x}{L\,sh\,x}\,L\,sh\,x\,dx =$$

$ch^2 - sh^2 = 1$

$$\frac{1}{2}L\int_{ch^{-1}[1]}^{ch^{-1}[\sqrt{2}]} ch^2 x\,dx = \frac{L}{4}\int_{ch^{-1}[1]}^{ch^{-1}[\sqrt{2}]} (1 + ch\,2x)\,dx =$$

$$= \frac{L}{4}\left[x + \frac{sh\,2x}{2} \right]_{ch^{-1}[1]}^{ch^{-1}[\sqrt{2}]} = \frac{L}{4}\Big(ch^{-1}[\sqrt{2}] - ch^{-1}[1] + \frac{1}{2}sh\big[2\,ch^{-1}[\sqrt{2}]\big] - \frac{1}{2}sh\big[2\,ch^{-1}[1]\big] \Big)$$

$$= \frac{L}{4}\left(ch^{-1}[\sqrt{2}] + \sqrt{2} \right) = \frac{L}{4}\left(\sqrt{2} + \log[1+\sqrt{2}] \right)$$

Donde hemos usado la relación: $sh[2\,ch^{-1}z] = 2z\,sh[ch^{-1}z] = 2z(z^2-1)^{\frac{1}{2}}$

Ahora recordamos el valor de ch^{-1}:

$$y = \frac{e^x + e^{-x}}{2} \Rightarrow 2ye^x = e^{2x} + 1 \Rightarrow e^{2x} - 2ye^x + 1 = 0$$

$$e^x = \frac{2y \pm (4y^2 - 4)^{\frac{1}{2}}}{2} = y \pm (y^2-1)^{\frac{1}{2}} \Rightarrow x = \log\left[y \pm (y^2-1)^{\frac{1}{2}} \right],\ y > 1$$

así que $ch^{-1}[\sqrt{2}] = \log\left[\sqrt{2} \pm (2-1)^{\frac{1}{2}}\right] = \log[\sqrt{2} \pm 1]$ tenemos que escoger la solución positiva $\Rightarrow ch^{-1}[\sqrt{2}] = \log[1+\sqrt{2}]$.

b/ El juego de los Euromillones consiste en elegir, por un lado, cinco números entre el 1 y el 49 y, por otro lado, dos números (llamados estrellas) entre el 1 y el 12. ¿Cuál es la probabilidad de acertar exactamente la combinación ganadora? ¿Y cuál es la probabilidad de acertar una de las estrellas y tres de los números?

i)

$$p = \frac{\text{nº de combinaciones ganadoras}}{\text{nº de combinaciones posibles}} = \frac{1}{C^{49}_{5} \times C^{12}_{2}} = \frac{1}{\frac{49!}{5!\,44!} \cdot \frac{12!}{2!\,10!}} =$$

$$= \frac{5!\,44!\;2!\,10!}{49!\;12!} = \frac{5\cdot\cancel{4\cdot3}\;\cancel{2\cdot2}}{49\cdot\cancel{48}\cdot47\cdot46\cdot45\;12\cdot11} = \frac{1}{49\cdot12\cdot47\cdot46\cdot9\cdot11} = \frac{1}{125854344} \sim 7.94\cdot10^{-9}$$

ii) Existen dos interpretaciones, ambas satisfactorias: o bien nos piden la probabilidad de acertar exactamente 3 números y exactamente 1 estrella o bien nos piden la probabilidad de acertar al menos 3 números y al menos 1 estrella

Probabilidad de acertar exactamente 3 números y 1 estrella:

$$\frac{\text{nº de ordenaciones con exactamente 3 números acertados} \times \text{nº de ordenaciones con exactamente 1 estrella acertada}}{\text{nº de ordenaciones totales}} = \frac{\binom{5}{3}\binom{44}{2}\binom{2}{1}\binom{10}{1}}{\binom{49}{5}\binom{12}{2}} = \boxed{\frac{2150}{1430163} \sim 0{,}00150}$$

Probabilidad de acertar al menos 3 números y al menos 1 estrella:

$$= \frac{\Bigg[\underbrace{\binom{5}{3}\binom{44}{2}}_{\substack{\text{Ordenaciones con exactamente}\\ \text{3 números acertados}}} + \underbrace{\binom{5}{4}\binom{45}{1}}_{\substack{\text{Ordenaciones con exactamente}\\ \text{4 números acertados}}} + \underbrace{\binom{5}{5}\binom{45}{0}}_{\substack{\text{Ordenaciones con}\\ \text{los 5 números acertados}}}\Bigg]\Bigg[\underbrace{\binom{2}{1}\times\binom{10}{1}}_{\substack{\text{Ordenaciones con exactamente}\\ \text{una estrella acertada}}} + \underbrace{\binom{2}{2}\binom{11}{0}}_{\substack{\text{Ordenaciones con las 2}\\ \text{estrellas acertadas}}}\Bigg]}{\binom{49}{5}\binom{12}{2}} = \boxed{\frac{4843}{2996532} \to 0.00162}$$

c/ Supongamos una máquina situada en $x = 0$ que lanza pelotas a una velocidad v_0 a lo largo de una mesa con rozamiento. La deceleración producida por el rozamiento es uniforme y, por lo tanto, la posición de la pelota está dada por $x = v_0 t - \frac{1}{2}at^2$. En cuanto una pelota se detiene, la máquina vuelve a lanzar otra pelota desde $x = 0$. Obtenga la distribución de probabilidad $f(x)dx$ de encontrar la pelota en el intervalo $(x, x + dx)$. Sugerencia: la distribución de probabilidad $f(x)dx$ es proporcional al tiempo dt que tarda la pelota en recorrer dicho intervalo.

Ecuación del movimiento: $x = v_0 t - \frac{1}{2}at^2 \Rightarrow$ Ecuación de la velocidad $v = v_0 - at$

Tiempo que tarda la pelota en detenerse $\Rightarrow T = \frac{v_0}{a}$

La distancia x_{max} recorrida por la pelota hasta detenerse será entonces:

$\Rightarrow x_{max} = \frac{v_0^2}{a} - \frac{1}{2}a\frac{v_0^2}{a^2} = \frac{v_0^2}{2a} \Rightarrow$ La nueva variable aleatoria x varía en el intervalo $\left[0, \frac{v_0^2}{2a}\right]$

De la ecuación del movimiento despejamos el valor de t:

Tenemos que escoger "$-$" porque para $x=0 \Rightarrow t=0$

$$t = \frac{v_0 \pm (v_0^2 - 2ax)^{1/2}}{a} = \frac{v_0}{a} - \frac{(v_0^2 - 2ax)^{1/2}}{a} \Rightarrow$$

de aquí se deduce que: $dt = \dfrac{dx}{(v_0^2 - 2ax)^{1/2}}$

Utilizamos ahora la sugerencia: $f(x)dx = k\,dt = \dfrac{k\,dx}{(v_0^2 - 2ax)^{1/2}}$

Para determinar la constante k usamos la condición de normalización y recordamos que $0 \leq x < \dfrac{v_0^2}{2a}$:

$$\int_0^{\frac{v_0^2}{2a}} \frac{k\,dx}{(v_0^2 - 2ax)^{1/2}} = k\left[\frac{(v_0^2 - 2ax)^{1/2}}{\frac{1}{2}(-2a)}\right]_0^{\frac{v_0^2}{2a}} = k\,\frac{v_0}{a} = 1 \rightarrow k = \frac{a}{v_0}$$

Por tanto el resultado final es

$$\boxed{f(x) = \frac{a}{v_0(v_0^2 - 2ax)^{1/2}} \quad 0 \leq x < \frac{v_0^2}{2a}}$$

Queremos comprobar si en un casino están utilizando barajas de cartas trucadas que tienen menos de cuatro ases por baraja (de cuarenta cartas), para disminuir la probabilidad de ganancia de los clientes. Para ello, observamos un crupier que está sacando cartas de una baraja. Después de sacar cada carta, la devuelve a la baraja y barajea. De N cartas que van apareciendo, contamos M ases. ¿Cuál es el número máximo de ases que hemos tenido que observar para asegurar que dicha baraja está trucada con un nivel central de confianza de X % (o alternativamente una confianza de $\gamma\sigma$, siendo γ un número positivo dado)? Suponga que $N \gg 1$.

Como devolvemos la carta que ha salido después de cada extracción tenemos

N experimentos aleatorios independientes → éxito sale as / → fracaso no sale as.

Distribución de población $P(x|p) = \binom{N}{x} p^x (1-p)^{N-x}$ (distribución binomial)

x = nº de ases obtenidos.

Si $N \gg 1$ aproximamos la distribución binomial por una normal.

$$P(x|p) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad \mu = Np \text{ y } \sigma = \sqrt{Np(1-p)}$$

Queremos estimar p. Para ello utilizamos el estimador $\hat{p} = \frac{x}{N}$.

La distribución muestral $P(\hat{p}|p)$ se obtiene mediante el cambio de variables aleatorias $x = N\hat{p}$ en la distribución de población:

$$p(\hat{p}|p) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(N\hat{p}-\mu)^2}{2\sigma^2}\right] N = \frac{1}{\sqrt{2\pi \frac{p(1-p)}{N}}} \exp\left[-\frac{N^2(\hat{p}-p)^2}{2N\, p(1-p)}\right] =$$

$$= \frac{1}{\sqrt{2\pi}\left(\frac{p(1-p)}{N}\right)^{\frac{1}{2}}} \exp\left[-\frac{(\hat{p}-p)^2}{2\frac{p(1-p)}{N}}\right] = \frac{1}{\hat{\sigma}\sqrt{2\pi}} \exp\left[-\frac{(\hat{p}-p)^2}{2\hat{\sigma}^2}\right]$$

La distribución muestral es una normal de media p y desviación típica $\hat{\sigma} = \left(\frac{p(1-p)}{N}\right)^{\frac{1}{2}}$.

El número máximo de ases M necesarios para asegurar que la baraja está trucada será aquél para el cual p_+, que es el límite superior del intervalo de confianza para p, vale $\frac{4}{40} = \frac{1}{10}$

$$p_+ = \frac{4}{40} = \frac{1}{10}.$$

El límite superior del intervalo de confianza se calcula con la fórmula

$$\int_{-\infty}^{\hat{p}_{obs}} P(\hat{p}|p_+)\, d\hat{p} = \alpha \qquad \text{donde } \hat{p}_{obs} = \frac{M}{N}$$

M = número máximo de ases observados para asegurar que la baraja está trucada

$$\int_{-\infty}^{\hat{p}_{obs}} \frac{1}{\hat{\sigma}\sqrt{2\pi}} \exp\left[-\frac{(\hat{p}-p_t)^2}{2\hat{\sigma}_t^2}\right] d\hat{p} = \left[\begin{array}{l} \frac{\hat{p}-p_t}{\hat{\sigma}} = z \\ dz = \frac{1}{\hat{\sigma}_t}\, d\hat{p} \end{array}\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\hat{p}_{obs}-p_t}{\hat{\sigma}_t}} e^{-\frac{z^2}{2}}\, dz =$$

$$= \Phi\left[\frac{\hat{p}_{obs}-p_t}{\hat{\sigma}_t}\right] = \alpha \Rightarrow \frac{\hat{p}_{obs}-p_t}{\hat{\sigma}_t} = \Phi^{-1}[\alpha] \Rightarrow \hat{p}_{obs} = \hat{\sigma}_t\, \Phi^{-1}[\alpha] + p_t$$

Sustituyendo $\hat{p}_{obs} = \frac{M}{N}$ y $p_t = \frac{1}{10}$ obtenemos: $\hat{\sigma}_t = \left(\frac{\frac{1}{10}\left(\frac{9}{10}\right)}{N}\right)^{\frac{1}{2}} = \frac{3}{10\sqrt{N}}$

$$\frac{M}{N} = \frac{1}{10} + \frac{3}{10\sqrt{N}}\, \Phi^{-1}[\alpha]$$

Nivel de confianza $\gamma\sigma \Rightarrow P_r(\mu-\gamma\sigma \le X \le \mu+\gamma\sigma) = 2\Phi(\gamma)-1$

$$1-2\alpha = 2\Phi(\gamma)-1 \Rightarrow 2\Phi(\gamma) = 2(1-\alpha) \Rightarrow \Phi(\gamma) = 1-\alpha \Rightarrow \alpha = 1-\Phi(\gamma) \Rightarrow$$

$$\Rightarrow \Phi^{-1}[\alpha] = \Phi^{-1}[1-\Phi(\gamma)] = -\Phi^{-1}[\Phi(\gamma)] = -\gamma$$

Por tanto, el resultado final en términos de γ es

$$\boxed{\frac{M}{N} = \frac{1}{10} - \frac{3\gamma}{10\sqrt{N}} \Rightarrow \qquad M = \frac{N}{10}\left(1-\frac{3\gamma}{\sqrt{N}}\right)}$$

Nivel central de confianza de $X\,\%$ $\quad 1-2\alpha = \frac{X}{100} \Rightarrow \alpha = \frac{1}{2}(1-\xi)$, $\xi = \frac{X}{100}$

Sustituyendo en la fórmula de arriba obtenemos:

$$\boxed{\frac{M}{N} = \frac{1}{10} + \frac{3}{10\sqrt{N}}\, \Phi^{-1}\left[\frac{1}{2}(1-\xi)\right]}$$