Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Matemática

Profesor Rodrigo Assar
Auxs: Sergio Araneda & Felipe Campos
Otoño 2012

MA3403-4: PROBABILIDADES Y ESTADÍSTICA
PAUTA CONTROL 2

1. Pregunta 1

   a) *A:* Sea X un v.a. y $x_1, \dots, x_n$ valores muestrales, i.e. valores obtenidos de realizaciones independientes $X_1, \dots, X_n$ de la misma variable X. Denote $E(X) = \mu$ y $Var(X) = \sigma^2$.

   1) *(0.6)* Explique las similitudes y diferencias entre la esperanza y el promedio. Use algún ejemplo ilustrativo.

      La esperanza es la versión del promedio para el caso probabilista. (0.2) A1.1
      Se trata de un promedio ponderado por la probabilidad de selección de cada valor posible. (0.2) A1.2
      Por Ejemplo si consideramos X el lanzamiento de una moneda codificando cara como 1 y sello con 0, suponiendo ambos resultados con igual probabilidad, la esperanza es el promedio 0.5. Pero si por ejemplo la cara tiene prob. $0{,}6$ entonces la esperanza es $0{,}6 \cdot 1 + 0{,}4 \cdot 0 = 0{,}6$. (0.2 ejemplo) A1.3

   2) *(0.7): Ley débil de los Grandes Números.* Muestre que $E(\bar{X}) = \mu$ con $\bar{X}$ el promedio de los X_i. Usando Tchebychev muestre que $\mathbb{P}(|\bar{X} - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$ y deduzca que $\bar{X}$ converge en probabilidad a μ.

      Consideremos Tchebychev para la variable $\bar{X}_n$.
      Se tiene que $\mathbb{E}(\bar{X}_n) = \mathbb{E}(\frac{\sum_{i=1}^n X_i}{n}) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n}n\mu = \mu$ (0.2) A2.1
      Así Tchebychev $\Rightarrow \mathbb{P}(|\bar{X}_n - \mu| \geq \epsilon) \leq \frac{Var(X)}{\epsilon^2}$ (0.2) A2.2
      Como $Var(\bar{X}_n) = Var(\frac{\sum_{i=1}^n X_i}{n}) = \frac{1}{n^2}Var(\sum_{i=1}^n X_i) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$ (0.3) A2.3
      Se concluye $\mathbb{P}(|\bar{X} - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$.

      - ¿El alumno comenta correctamente la definición de convergencia en probabilidad con lo obtenido? A2.4

   3) *(0.7)* Sea $S_n^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{X})^2$, muestre que $E(S_n^2) = \frac{n-1}{n}\sigma^2$. Sea $\tilde{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{X})^2$, muestre que $E(\tilde{S}_n^2) = \sigma^2$. Compare la convergencia de $E()$ para ambos estimadores de la varianza, cuál preferiría y por qué?

      $\mathbb{E}(S_n^2) = \mathbb{E}(\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2)$
      $= \frac{1}{n}\mathbb{E}(\sum_{i=1}^n (X_i - \bar{X}_n)^2) = \frac{1}{n}\mathbb{E}(\sum_{i=1}^n X_i^2 - \mu^2 + \bar{X}_n^2 + \mu^2 - 2X_i\bar{X}_n)$
      $= \frac{1}{n}(\sum_{i=1}^n (\mathbb{E}(X_i^2) - \mu^2) + \mathbb{E}(\bar{X}_n^2) + \mu^2 - 2\mathbb{E}(X_i\bar{X}_n))$

$= \frac{1}{n}(\sum_{i=1}^{n} Var(X) + \mathbb{E}(\bar{X}_n^2) + \mu^2 - 2\mathbb{E}(X_i\bar{X}_n))$
$= \frac{1}{n}(nVar(X) + n\mathbb{E}(\bar{X}_n^2) + n\mu^2 - 2\mathbb{E}(\sum_{i=1}^{n} X_i\bar{X}_n))$
$= \frac{1}{n}(nVar(X) + n\mathbb{E}(\bar{X}_n^2) + n\mu^2 - 2\mathbb{E}(n\bar{X}_n^2))$
$= \frac{1}{n}(nVar(X) + n(\mu^2 - \mathbb{E}(\bar{X}_n^2)))$
$= \frac{1}{n}(nVar(X) - nVar(\bar{X}_n))) = \frac{1}{n}(n\sigma^2 - n\sigma^2/n)) = \frac{n-1}{n}\sigma^2$ (0.3) A3.1

Entonces $\mathbb{E}(\tilde{S}_n^2) = \mathbb{E}(\frac{n-1}{n}S_n^2) = \frac{n-1}{n}\mathbb{E}(S_n^2) = \frac{n-1}{n}\frac{n}{n-1}\sigma^2 = \sigma^2$ (0.2) A3.2
En ambos casos se tiene que si $n \to \infty$, $\mathbb{E}(\tilde{S}_n^2)$ y $\mathbb{E}(S_n^2) \to \sigma^2$. Pero $\mathbb{E}(\tilde{S}_n^2)$ es exactamente σ^2 al menos con este criterio preferiría $\tilde{S}_n^2$ que converge más rápido. (0.2) A3.3

- ¿El alumno sabe que $Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$? A3.4
- ¿El alumno elige un criterio y argumenta coherentemente respecto a el para elegir entre S y S tilda? A3.5
- ¿El alumno elige un criterio útil? ¿Cuál? A3.6

b) *B: Teorema del mono infinito.* Pongamos a un mono a escribir en el teclado del computador (ver figura) y veamos qué pasa. Suponga que el teclado tiene 50 teclas (bloqueemos las otras) y la palabra que se quiere escribir es *banana*. Asumamos que cada tecla es presionada aleatoriamente, con igual probabilidad e independientemente.

1) *(0.6)*Calcule la probabilidad de que las primeras 6 letras que el mono escriba sean *banana*.

$(\frac{1}{50})^6$. El exponente 6 corresponde a las 6 letras asociadas a la palabra banana y la multiplicación gracias a la independencia entre tecleos. (0.3) B1.1
$1/50$ Es la probabilidad de escoger una letra particular del teclado (0.3) B1.2

- ¿El alumno comprendió el enunciado? B1.3

2) *(0.7)*Suponga que ahora tenemos a 100 monos escribiendo cada uno frente a un computador distinto e independientemente. Calcule la probabilidad de que ningún mono logre escribir *banana* al primer intento (6 primeras letras).

$\mathbb{P}$(Ningun acierto entre 100) $= (1 - (1/50^6))^{100}$. El exponente 100 es por la independencia y la cantidad de monos. (0.3) B2.1
$(1 - (1/50^6))$ Es la probabilidad de no escribir banana (parte anterior) (0.4) B2.2

3) *(0.7)*Deduzca que si la cantidad de monos tiende a ∞ entonces la probabilidad de que ningún mono acierte converge a 0.

Una generalización del argumento anterior implica que $\mathbb{P}$(Ningun acierto entre n) $=$

$(1-(1/50^6))^n$. (0.3) B3.1
$(1-(1/50^6)) < 1$, luego al hacer tender n a infinito la potencia $(1-(1/50^6))^n$ converge a 0. (0.4) B3.2

c) *C:* Considere una fábrica de lápices en que la probabilidad de que un lápiz cualquiera resulte fallado es p , un parámetro desconocido. Para estimar el valor de p el dueño ha dispuesto la siguiente experiencia:
Durante el día se escogerán 20 lápices de los fabricados y se contará la cantidad de fallados. Luego se estimará p como la cantidad de fallados dividida 20.

1) *(0.6)*Añada requerimientos para que la cantidad de fallados (X) siga una binomial de parámetros 20 y p.

   Se desea que X: Cantidad fallados siga una distribución binomial $(20, p)$. Para que sea binomial cada sub experimento (sacar un lápiz) debe corresponder a una v.a. bernoulli y ser independiente de las demás. (0.3) C1.1
   Para esto se requiere que los lápices sean sacados CON REPOSICIÓN. Es decir, se saca un lápiz, se devuelve, y así hasta sacar 20. (0.3) C1.2
   Extra: Si el número de lápices totales $N >> 20$. La variable X: cantidad de fallados (que en principio es una hipergeométrica) puede ser aproximada por una binomial. C1.3

2) *(0.7)*Con dichos requerimientos añadidos, demuestre que $E(X) = np$ y $Var(X) = np(1-p)$. Deduzca que $E(\bar{X}) = p$ y que $Var(\bar{X})$ converge a 0 cuando n converge a ∞. Qué quieren decir estos resultados respecto a $\bar{X}$ y p?.

   Con f.g.m.
   $\mathbb{E}(e^{tX}) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{tk}$
   $= \sum_{k=0}^{n} \binom{n}{k} (e^t p)^k (1-p)^{n-k} = (pe^t + 1 - p)^n$

   $\mathbb{E}(X^k) = M_X^{(k)}(0)$

   $\frac{d}{dt}(pe^t + 1 - p)^n = n(pe^t + 1 - p)^{n-1} pe^t$
   $\mathbb{E}(X) = M_X'(0) = n(p + 1 - p)^{n-1} p = n \cdot 1 \cdot p = np$ (0.2) C2.1

   $\frac{d^2}{dt^2}(pe^t + 1 - p)^n = npe^t(n-1)(pe^t + 1 - p)^{n-2} pe^t + n(pe^t + 1 - p)^{n-1} pe^t$
   $\mathbb{E}(X^2) = M_X''(0) = np(n-1)p + np = n^2p^2 - np^2 + np = np(1-p) + n^2p^2$
   $Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = np(1-p) + n^2p^2 - (np)^2 = np(1-p)$ (0,2) C2.2

   (Si el alumno no ocupa f.g.m corresponde el puntaje al calculo de la esperanza y varianza de igual forma)

   Definiendo X_i: el lápiz i sacado está fallado. X_i tiene distribución bernoulli(p), luego

$\mathbb{E}(X_i) = 1 \cdot p + 0 \cdot (1-p) = p$ y por lo tanto $\mathbb{E}(\bar{X}) = \mathbb{E}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{1}{n}np = p$. (0,1) C2.3

Además $Var(\bar{X}) = Var(\frac{1}{n}\sum_{i=1}^{n} X_i) = Var(\frac{X}{n}) = \frac{1}{n^2}Var(X) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n}$.
$\lim_{n\to\infty} \dfrac{p(1-p)}{n} = 0$ (0,1) C2.4

Esto quiere decir que $\bar{X}$ se aproxima a p para n muy grande. Es decir, con n grande $\bar{X}$ es un buen estimador de p. (0,1) C2.5

- ¿El alumno usa y define correctamente la función generadora de momentos? C2.6

3) *(0.7)*Compare el histograma (gráfico de frecuencias) asociado a $\bar{X}$ con la gráfica de la densidad de una normal. Cuál es el rol de la densidad para v.as continuas en comparación con la función de probabilidad en v.as discretas?

   (Ver pauta a mano) Podemos ver que al tender N a infinito, la distribución binomial adquiere forma de campana similar a la función de densidad de una v.a. gaussiana. (0.4 gráfico similitud) C3.1

   La idea de una función de densidad es que $f_X(x_0)\Delta x_0 \approx \mathbb{P}(x \in [x_0 - \Delta x_0, x_0 + \Delta x_0])$ Esto para $\Delta x_o = dx_0 \to 0$ permite ver que para una v.a. continua con densidad

$$\mathbb{P}(a \le X \le b) = \int_a^b f_X(x)dx$$

   (0,3) C3.2

   - ¿El alumno muestra algún ejemplo correcto para comparar la analogía discreta de una función de densidad? C3.3

d) Pregunta 2

1) *A:* Sea $X \sim Bin(10, \frac{1}{2})$, es decir $\mathbb{P}(X = k) = \binom{10}{k} 0{,}5^{10}$ $\forall k = 0, \ldots, 10$.

   a′ *(1.0)*Calcule de manera exacta $\mathbb{P}(X \ge 7)$.

   $\mathbb{P}(X \ge 7) = \mathbb{P}(X = 7) + \mathbb{P}(X = 8) + \mathbb{P}(X = 9) + \mathbb{P}(X = 10)$ (0.5) D1.1
   $= (1/2)^{10}(\binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}) = 0{,}1718$ (0.5) D1.2

   - ¿El alumno conoce $\mathbb{P}(X = k)$ para una v.a. binomial? D1.3

b' *(1.0)*Usando el TCL (Teorema Central del Límite) aproxime $\mathbb{P}(X \geq 7)$ por 0,1038.

TCL :

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \rightarrow^{D} \mathcal{N}(0,1)$$

(0.1) D2.1

Aquí cada X_i es una v.a. bernoulli(1/2).

Luego $\mathbb{P}(X \geq 7) = \mathbb{P}(\dfrac{X - 10 \cdot 1/2}{\sqrt{10 \cdot 1/2 \cdot 1/2}} \geq \dfrac{7 - 10 \cdot 1/2}{\sqrt{10 \cdot 1/2 \cdot 1/2}})$ (0,6) D2.2

$\approx \mathbb{P}(Z_n \geq 1{,}26) \approx 0{,}1038$ Donde Z_n es una v.a. normal$(0,1)$ (0.3) D2.3

- ¿El alumno reconoce correctamente la varianza de la binomial? D2.4
- ¿El alumno reconoce que una v.a. binomial se puede ver como el promedio de v.a. Bernoulli? D2.5

c' *(1.0)*Compare ambos resultados e interprete.

Ambos números difieren en 0.068

$$0{,}1718 - 0{,}1038 = 0{,}068$$

(0.5) D3.1
$n = 10$ no es muy grande. (0.5) D3.2

2) *B:* Un hombre y una mujer deciden encontrarse en un café. Se tiene que cada persona llega uniformemente distribuida entre 12:00 y 13:00 hrs.

a' *(1.0)*Encuentre la probabilidad de que el primero en llegar no tenga que esperar al segundo más de 10 minutos.

Sea H: Tiempo de llegada del hombre. M: Tiempo llegada mujer. Por el enunciado, H y M son v.a. Uniforme $(0,60)$ (0,2) E1.1

$\mathbb{P}(|H - M| \leq 10) = \mathbb{P}(-10 \leq H - M \leq 10) = \mathbb{P}(M - 10 \leq H \leq M + 10) = \int\int_R f_{HM} dhdm$, donde R es la región correspondiente al evento. (0,2) E1.2

Por la independencia de las llegadas, se tiene que $f_H M = f_H \cdot f_M = (1/60)^2 1_{[0,60]\times[0,60]}$ (0,2) E1.3

Como hay una constante, esta sale afuera de la integral y basta con calcular el área de $R \cap [0,60] \times [0,60]$. (Ver pauta a mano diagrama)

Area $= (60 \cdot 60) - (50 \cdot 50)/2 = 60^2 - 50^2$
$\Rightarrow \mathbb{P}(M - 10 \leq H \leq M + 10) = \frac{1}{60^2}(60^2 - 50^2) = 1 - (5/6)^2 = 11/36$ (0,4) E1.4

- ¿El alumno intenta calcular el área de la integral usando un método distinto al geométrico usado en la pauta? E1.5

b' *(1.0)*Encuentre la probabilidad de que la suma de la cantidad de minutos después de las 12:00 hrs. en que llegan sea menor a 30.

$\mathbb{P}(H + M < 30) = \mathbb{P}(H < 30 - M)$ (0,4) E2.1

Dibujar gráfico. Ver pauta escrita diagrama área. (0,3) E2.2

$\mathbb{P}(H + M < 30) = \frac{30 \cdot 30}{2} \frac{1}{60 \cdot 60} = \frac{1}{8}$
$\Rightarrow \mathbb{P}(H + M < 30) = \frac{1}{8}$ (0,3) E2.3.

Se puede hacer lo mismo sin gráfico, si lo hacen tomar el puntaje del gráfico en vez.

- ¿El alumno intenta calcular el área de la integral usando un método distinto al geométrico usado en la pauta? E2.4

3) C *(1.0):* Suponga que V, la rapidez de un objeto, tiene distribución $N(0, 4)$. Si $K = \frac{MV^2}{2}$ es la energía cinética del objeto, encuentre la densidad de K. Si la masa $M = 10$, calcule $\mathbb{P}(K \leq 3)$ y $E(K)$.

V es $\mathcal{N}(0, 4)$. Densidad de V: $f_V(v) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{v}{2})^2} = \frac{1}{2\sqrt{2\pi}} e^{-\frac{v^2}{8}}$
$F_K(k) = \mathbb{P}(K \leq k) = \mathbb{P}(\frac{1}{2} m V^2 \leq k)$
$= \mathbb{P}(V^2 \leq \frac{2k}{m}) = \mathbb{P}(|V| \leq \sqrt{\frac{2k}{m}}) = \mathbb{P}(-\sqrt{\frac{2k}{m}} \leq V \leq \sqrt{\frac{2k}{m}})$
$= \mathbb{P}(V \leq \sqrt{\frac{2k}{m}}) - \mathbb{P}(V \leq -\sqrt{\frac{2k}{m}}) = F_V(\sqrt{\frac{2k}{m}}) - F_V(-\sqrt{\frac{2k}{m}})$ (0,2) F1.1

$F_K'(k) = f_K(k) = (\sqrt{\frac{2k}{m}})' \cdot f_V(\sqrt{\frac{2k}{m}}) - (-\sqrt{\frac{2k}{m}})' \cdot f_V(-\sqrt{\frac{2k}{m}})$ (0,1) F1.2
$= \sqrt{\frac{2}{m}} \frac{1}{2} k^{-\frac{1}{2}} \frac{1}{2\sqrt{2\pi}} e^{-\frac{k}{4m}} + \sqrt{\frac{2}{m}} \frac{1}{2} k^{-\frac{1}{2}} \frac{1}{2\sqrt{2\pi}} e^{-\frac{k}{4m}} = \sqrt{\frac{2}{m}} k^{-\frac{1}{2}} \frac{1}{2\sqrt{2\pi}} e^{-\frac{k}{4m}} = \sqrt{\frac{1}{4\pi m}} k^{-\frac{1}{2}} e^{-\frac{k}{4m}}$

$\Rightarrow \forall k \geq 0 \quad f_K(k) = \sqrt{\frac{1}{4\pi m}} k^{-\frac{1}{2}} e^{-\frac{k}{4m}}$ (0,2) F1.3

$m = 10$

$$f_K(k) = \sqrt{\frac{1}{40\pi}} k^{-\frac{1}{2}} e^{-\frac{k}{40}}$$

Por lo anterior $\mathbb{P}(K \leq 3) = F_V(\sqrt{\frac{2\cdot 3}{10}}) - F_V(-\sqrt{\frac{2\cdot 3}{10}})$.
Normalizando V se puede expresar la probabilidad pedida en base a probabilidades de una normal(0,1).
$\mathbb{P}(V \leq \sqrt{\frac{3}{5}}) - \mathbb{P}(V \leq -\sqrt{\frac{3}{5}})$
$\mathbb{P}(\frac{V}{2} \leq \sqrt{\frac{3}{5}}) - \mathbb{P}(\frac{V}{2} \leq -\sqrt{\frac{3}{5}}) = 1 - 2\mathbb{P}(Z > 0,39) = 1 - 2\cdot 0,3438 = 0,3124$ (0,3) F1.4

$\mathbb{E}(\frac{MV^2}{2}) = \frac{M}{2}\mathbb{E}(V^2) = \frac{M}{2}(\sigma^2 + \mathbb{E}(V)^2) = \frac{10}{2}(4+0) = 20$ (0,2) F1.5

e) Pregunta 3

1) *A:* Suponga que los artículos de Wikipedia tienen una extensión media de 1000 caracteres (letras) y una desviación estandard de 200 caracteres.

   a' *(1.0)*Dé, usando Tchebychev, una cota inferior para la probabilidad de encontrar artículos con extensión entre 600 y 1400 caracteres.

   De Tchebychev $\mathbb{P}(|X - 1000| \geq 400) \leq \frac{200^2}{400^2} = \frac{1}{4}$ (0,5) G1.1

   Por lo tanto $\mathbb{P}(|X - 1000| \leq 400) \geq \frac{3}{4}$ (0,5) G1.2

   - ¿El alumno llegó a modelar el evento pedido como $\{\omega : |X(\omega) - 1000| \geq 400\}$? G1.3

   b' *(1.0)*Usando la desigualdad de Markov directamente, pruebe que $\mathbb{P}(X \geq 600) \leq 1$ y que $\mathbb{P}(X \leq 1400) \geq \frac{2}{7}$, donde X es la v.a. extensión.

   Usando markov con $\epsilon = 600$, $\mathbb{P}(X \geq 600) \leq \frac{1000}{600} = \frac{5}{3}$, Pero como $\mathbb{P}$ es probabilidad, a lo más puede ser 1.(0,5) G2.1
   Es decir, la cota dada por markov no aporta información.

   Con $a = 400$, $\mathbb{P}(X \geq 1400) \leq \frac{1000}{1400} = \frac{5}{7}$, (0,5) G2.2
   $\Rightarrow \mathbb{P}(X \leq 1400) \geq \frac{2}{7}$.

   - ¿El alumno enuncia correctamente la desigualdad de markov? G2.3

   c' *(1.0)*Compare los resultados de las dos partes anteriores, cuál aporta más? Explique.

La cota deducida en i) implica que $\mathbb{P}(X \in [600, 1400]) \geq \frac{3}{4}$. Las cotas encontrada en ii) solo permiten concluir que $\mathbb{P}(X \in [600, 1400] \geq \frac{2}{7})$, una cota más baja ya que $\frac{3}{4} > \frac{2}{7}$ (0,5) G3.1

Luego la cota de la desigualdad de Tchebychev aporta más. (0,5) G3.2

2) *Aproximación de Binomial por Normal*
El teorema límite de DeMoivre-Laplace dice que, para z_1 y z_2 números fijos, una variable bernoulli B_N de parámetros (N, p) satisface

$$\lim_{n\to\infty} \mathbb{P}(Np + z_1\sqrt{Npq} \leq B_N \leq Np + z_2\sqrt{Npq}) = \Phi(z_1) - \Phi(z_2)$$

Una fracción desconocida p de cierta población son fumadores. Se desea conocer p escogiendo una muestra aleatoria de tamaño N con remplazo, y se desea que el error en la estimación no exceda a 0,001.

a' Justifique por qué en este modelo de muestreo la variable aleatoria que cuenta la cantidad de fumadores en la muestra sigue una distribución binomial (N,p).

Al hacer el muestreo, la probabilidad de sacar 1 fumador es justamente la fracción real p de fumadores de la población (bernoulli). Al haber reposición, la fracción de fumadores no va cambiando a medida que se sacan individuos.(0,8) H1.1
Por lo tanto la cantidad de fumadores en una muestra N es la cantidad de v.a. bernoullis i.i.d que resultan cara. Luego corresponde a una v.a. binomial de parámetros N y p. (0,7) H1.2

b' Usando la aproximación de DeMoivre-Laplace, estime N suficiente para obtener la cota del error deseada con probabilidad 0,95.

Sea p' la fracción de fumadores obtenidos en la muestra. Se pide entonces que $\mathbb{P}(|p' - p| < 0{,}001) \geq 0{,}95$. $p'N$ es el número de fumadores obtenidos al sacar la muestra, denotemos F esta v.a. .Por la parte a) F sigue una distribución binomial (N, p). Multiplicando por N se obtiene $\mathbb{P}(|Np' - Np| < 0{,}001N) = \mathbb{P}(|F - Np| < 0{,}001N)$. Esto es equivalente a

$$\mathbb{P}(Np - 0{,}001N < F < NP + 0{,}001N) \geq 0{,}95$$

(0,5) H2.1

Entonces definiendo z tal que $z\sqrt{Npq} = 0{,}001N$, $z = \frac{0{,}001\sqrt{N}}{\sqrt{pq}}$, se puede usar el teorema de DeMoivre-Laplace. Para obtener una probabilidad mayor a 0.95, basta encontrar z en la tabla de distribución tal que $\Phi(z) - \Phi(-z) = 0{,}95$. (Definir z adecuado 0,5) H2.2

Buscando en la tabla, con $z \geq 1{,}96$ se tiene la probabilidad mayor a $0{,}95$. Luego $\frac{0{,}001\sqrt{N}}{\sqrt{pq}} \geq 1{,}96 \Rightarrow N \geq 1{,}96^2 pq/0{,}001^2$. Usando la pista para acotar $pq = p(1-p)$, basta tomar N tal que $N \geq \frac{1}{4}\frac{1{,}96^2}{0{,}001^2} = 960400$ para satisfacer la condición sobre z y así la probabilidad pedida. (0,5) H2.3

Para tener una cota del error tan fina como $0{,}001$, es necesaria una muestra enorme.

Hint: $p(1-p) \leq 1/4 \quad p \in [0,1]$
Hint2: Muestreo con remplazo significa que de una población de M individuos, se eligen N individuos uno a la vez, se revisa si el individuo es fumador o no, y el individuo elegido vuelve a ser parte de la población escogible en las elecciones siguientes.

General:

- ¿El alumno explica en forma ordenada sus resultados? L.1
- ¿El alumno Entrega una hoja limpia? L.2
- ¿La letra del alumno es legible? L.3