Universidad Nacional Autónoma de México

Probabilidad I

Tarea Examen

Alan Ernesto Arteaga Vázquez

Raúl Llamosas Alvarado

Edgar Quiroz Castañeda

Jean Paul Ruiz Melo

Sandra Del Mar Soto Corderi

Martes 4 de Diciembre del 2018

1. Considere el experimento de lanzar dos tetraedros cuyos lados están numerados del 1 al 4. Sean Y_1, Y_2 los números más pequeño y más grande obtenidos en las caras superiores respectivamente. Tenemos lo siguiente:

$\frac{y}{x}$	1	2	3	4
1	$(1,1)$	$(1,2)$	$(1,3)$	$(1,4)$
2	$(1,2)$	$(2,2)$	$(2,3)$	$(2,4)$
3	$(1,3)$	$(2,3)$	$(3,3)$	$(3,4)$
4	$(1,4)$	$(2,4)$	$(3,4)$	$(4,4)$

(a) Encuentre la función de densidad conjunta de Y_1, Y_2
De esto se tiene:

$\frac{y}{x}$	1	2	3	4	$f_X(x)$
1	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{7}{16}$
2	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{5}{16}$
3	0	0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$
4	0	0	0	$\frac{1}{16}$	$\frac{1}{16}$
$f_Y(y)$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{7}{16}$	

(b) Obtenga la densidad condicional de Y_2 dado Y_1 para cada uno de los posibles valores de Y_1

$$f(1|X) = \frac{P(X=x, Y=1)}{P(X=x)} = \frac{\frac{1}{16}}{\frac{7}{16}} + \frac{\frac{2}{16}}{\frac{7}{16}} + \frac{\frac{2}{16}}{\frac{7}{16}} + \frac{\frac{2}{16}}{\frac{7}{16}} = .14 + .28 + .28 + .28$$

$$f(2|X) = \frac{P(X=x, Y=2)}{P(X=x)} = \frac{0}{\frac{5}{16}} + \frac{\frac{1}{16}}{\frac{5}{16}} + \frac{\frac{2}{16}}{\frac{5}{16}} + \frac{\frac{2}{16}}{\frac{5}{16}} = .2 + .4 + .4$$

$$f(3|X) = \frac{P(X=x, Y=3)}{P(X=x)} = \frac{0}{\frac{3}{16}} + \frac{0}{\frac{3}{16}} + \frac{\frac{1}{16}}{\frac{3}{16}} + \frac{\frac{2}{16}}{\frac{3}{16}} = .33 + .66$$

$$f(4|X) = \frac{P(X=x, Y=4)}{P(X=x)} = \frac{0}{\frac{1}{16}} + \frac{0}{\frac{1}{16}} + \frac{0}{\frac{1}{16}} + \frac{\frac{1}{16}}{\frac{1}{16}} = 1$$

(c) Encuentre $E(Y_1Y_2), E(Y_1), E(Y_2)$

$$E(Y_1Y_2) = E(XY) = \frac{1}{16} + 2\frac{2}{16} + 3\frac{2}{16} + 4\frac{2}{16} + 0 + (2\times 2)\frac{1}{16} + (2\times 3)\frac{2}{16} + (2\times 4)\frac{2}{16} + 0 + 0 + (3\times 3)\frac{1}{16} + (3\times 4)\frac{2}{16} + 0 + 0 + 0 + (4\times 4)\frac{1}{16}$$

$$= \frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{8}{16} + \frac{4}{16} + \frac{12}{16} + \frac{16}{16} + \frac{9}{16} + \frac{24}{16} + \frac{16}{16} = \frac{84}{16} = 5.25$$

$$E(Y_1) = E(X) = \frac{7}{16} + 2(\frac{5}{16}) + 3(\frac{3}{16}) + 4(\frac{1}{16}) = 1.875$$

$$E(Y_2) = E(Y) = \frac{1}{16} + 2(\frac{3}{16}) + 3(\frac{5}{16}) + 4(\frac{7}{16}) = 3.125$$

(d) ¿Son Y_1 y Y_2 independientes?
Tenemos que para que sea independiente se debe cumplir que:

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

Entonces si $x = 1, y = 1$, ya vimos que $f_{X,Y}(1,1) = \frac{1}{16}$, $f_X(1) = \frac{7}{16}$ y $f_Y(1) = \frac{1}{16}$, entonces:

$$f_{X,Y}(1,1) = \frac{1}{16} \neq \frac{7}{256} = \frac{7}{16}\frac{1}{16} = f_X(1)f_Y(1)$$

Por lo tanto Y_1 y Y_2 no son independientes.

2. Suponga que X es una variable aleatoria con distribución $Bin(100, \frac{1}{5})$. Utilizando el teoerema de DeMoivre- Laplace, calcular $P(15 \leq X \leq 25)$.
Tenemos que DeMoivre-Laplace dice:

$$X - Bin(n,p) = \frac{x - np}{\sqrt{np(1-p)}} = z = \frac{x - 100(\frac{1}{5})}{\sqrt{100(\frac{1}{5})(\frac{4}{5})}} = \frac{x - 20}{4}$$

Para X-N(0,1), entonces:

$$P(15 \leq X \leq 25) = \phi(\frac{25-20}{4}) - \phi(\frac{15-20}{4}) = \phi(\frac{5}{4}) - \phi(\frac{-5}{4}) = \phi(\frac{5}{4}) - (1 - \phi(\frac{5}{4}))$$

$$\phi(1.25) - 1 + \phi(1.25) = 2() - 1 = 2(.8944) - 1 = 1.7888 - 1 = .7888$$

3. En una tarde soleada, Augusto lanza dos dados 2160 veces (él casi no tiene nada que hacer). Sea X el número de veces que 2 aparece en alguna de las caras. Encontrar la probabilidad de que X sea menor a 55.
Esto describe una Binomial con parámetros 2160 y $p = \frac{11}{36}$ entonces se sigue que

$$P(X \leq 54) = \sum_{i=0}^{54} \binom{2160}{i} (\frac{11}{36})^i (\frac{25}{36})^{2160-i}$$

Pero tomando en cuenta que la binomial son 2160 eventos bernoulli independientes se sigue que $\{X_i\}_{i=1}^{2160}$ son todos eventos bernoulli de forma que su esperanza y varianza son $Var(X) = \frac{(11)(25)}{36^2}$ y $E(X) = \frac{11}{36}$ y el teorema de limite central nos dice que

$$P\{X_1 + ...X_{2160} \leq 54\} = P\{\frac{\sum_{i=1}^{2160} X_i - 2160(\frac{11}{36})}{\sqrt{2160(\frac{11(25)}{36^2})}} \leq \frac{54 - 2160(\frac{11}{36})}{\sqrt{2160(\frac{11(25)}{36^2})}}\} \to \Phi(\frac{54 - 2160(\frac{11}{36})}{\sqrt{2160(\frac{11(25)}{36^2})}})$$

Y como

$$\frac{54 - 2160(\frac{11}{36})}{\sqrt{2160(\frac{11(25)}{36^2})}} = -28.306$$

Y esa probabilidad tiende a cero. Entonces

$$P(X < 55) \approx 0$$

4. Sean $X_1, ..., X_n$ variables aleatorias independientes. Use la función generadora de momentos para encontrar la distribución de $\sum_{i=1}^n X_i$ en los siguientes casos:

   (a) Si $\forall i \ \in \{1, .., n\}$ $X_i \sim Gamma(r, \lambda)$ Tenemos que si $Z = \sum_{i=1}^n X_i$ entonces la función generadora de momentos de Z denotada por $Mg_Z(t)$ está dada por:

   $$Mg_Z(t) = \prod_{i=1}^n Mg_{X_i}(t)$$

   pero como $\forall \ i \ Mg_{X_i}(t) = (\frac{\lambda}{\lambda - t})^r$ entonces

   $$\prod_{i=1}^n (\frac{\lambda}{\lambda - t})^r = (\frac{\lambda}{\lambda - t})^{\sum_{i=1}^n r} = (\frac{\lambda}{\lambda - t})^{nr}$$

   Pero si esa es la función generadora de Z entonces se tiene que

   $$Z \sim Gamma(nr, \lambda) \blacksquare$$

   (b) Si $\forall i \ \in \{1, .., n\}$ $X_i \sim Gamma(r_i, \lambda)$
   Análogamente como en el anterior, usamos la generadora de momentos y tenemos que

   $$Mg_Z(t) = \prod_{i=1}^n Mg_{X_i}(t) = \prod_{i=1}^n (\frac{\lambda}{\lambda - t})^{r_i} = (\frac{\lambda}{\lambda - t})^{\sum_{i=1}^n r_i}$$

   Entonces se sigue que

   $$Z \sim Gamma(\sum_{i=1}^n r_i, \lambda) \blacksquare$$

   (c) Si $\forall i \ \in \{1, .., n\}$ $X_i \sim exp(\lambda)$
   La función generadora de $Z = \sum_{i=1}^n X_i$, donde cada X_i tiene generadora de momentos $Mg_{X_i}(t) = \frac{\lambda}{\lambda - t}$ es

   $$Mg_Z(t) = \prod_{i=1}^n (\frac{\lambda}{\lambda - t}) = (\frac{\lambda}{\lambda - t})^n$$

   Pero esa es la generadora de momentos de una Gamma con parametro n, λ entonces:

   $$Z \sim Gamma(n, \lambda) \blacksquare$$

   (d) Si $\forall i \ \in \{1, .., n\}$ $X_i \sim Geo(p)$
   La función generadora de momentos de una variable aleatoria $Z = \sum_{i=1}^n X_i$ donde cada X_i tiene generadora de momentos $Mg_{X_i}(t) = \frac{p}{1-(1-p)e^t}$ está dada por

   $$Mg_Z(t) = \prod_{i=1}^n (\frac{p}{1 - (1 - p)e^t}) = (\frac{p}{1 - (1 - p)e^t})^n$$

   Lo cual es idéntica a una distribución binomial negativa con parámetro n,p.Entonces

   $$Z \sim BinNeg(n, p) \blacksquare$$

(e) Si $\forall i \ \in \{1,..,n\}$ $X_i \sim BinNeg(r,p)$
La función generadora de momentos de una variable aleatoria $Z = \sum_{i=1}^{n} X_i$ donde cada X_i tiene generadora de momentos $Mg_{X_i}(t) = (\frac{p}{1-(1-p)e^t})^r$ está dada por:

$$Mg_Z(t) = \prod_{i=1}^{n} (\frac{p}{1-(1-p)e^t})^r = (\frac{p}{1-(1-p)e^t})^{\sum_{i=1}^{n} r} = (\frac{p}{1-(1-p)e^t})^{nr}$$

Entonces se sigue que

$$Z \sim BinNeg(nr,p)\blacksquare$$

(f) Si $\forall i \ \in \{1,..,n\}$ $X_i \sim BinNeg(r_i,p)$
La función generadora de momentos de una variable aleatoria $Z = \sum_{i=1}^{n} X_i$ donde cada X_i tiene generadora de momentos $Mg_{X_i}(t) = (\frac{p}{1-(1-p)e^t})^{r_i}$ está dada por:

$$Mg_Z(t) = \prod_{i=1}^{n} (\frac{p}{1-(1-p)e^t})^{r_i} = (\frac{p}{1-(1-p)e^t})^{\sum_{i=1}^{n} r_i}$$

Entonces se sigue que

$$Z \sim BinNeg(\sum_{i=1}^{n} r_i, p)\blacksquare$$

(g) Si $\forall i \ \in \{1,..,n\}$ $X_i \sim Poisson(\lambda)$
La función generadora de momentos de una variable aleatoria $Z = \sum_{i=1}^{n} X_i$ donde cada X_i tiene generadora de momentos $Mg_{X_i}(t) = e^{\lambda(e^t-1)}$ está dada por:

$$Mg_Z(t) = \prod_{i=1}^{n} (e^{\lambda(e^t-1)}) = (e^{\lambda(e^t-1)})^n = (e^{n\lambda(e^t-1)})$$

Entonces se sigue que

$$Z \sim Poisson(n\lambda)\blacksquare$$

(h) Si $\forall i \ \in \{1,..,n\}$ $X_i \sim Bin(n,p)$
La función generadora de momentos de una variable aleatoria $Z = \sum_{i=1}^{n} X_i$ donde cada X_i tiene una generadora de momentos $Mg_{X_i}(t) = (e^t p + (1-p))^n$ está dada por:

$$Mg_Z(t) = \prod_{i=1}^{n} (e^t p + (1-p))^n = (e^t p + (1-p))^{n^2}$$

Entonces se sigue que

$$Z \sim Bin(n^2,p)\blacksquare$$

Nota: La notación del ejercicio infiere que el número de variables aleatorias coincide con el parámetro n. De ser éste el caso entonces distribuye como mencionamos arriba. Si ésta no era la intención entonces distribuye $Bin(n(k),p)$ donde k es el parámetro con el cual cada X_i distribuye de forma binomial.

(i) Si $\forall i \ \in \{1,..,n\}$ $X_i \sim Bin(n_i,p)$
La función generadora de momentos de una variable aleatoria $Z = \sum_{i=1}^{n} X_i$ donde cada X_i tiene una generadora de momentos $Mg_{X_i}(t) = (e^t p + (1-p))^{n_i}$ está dada por:

$$Mg_Z(t) = \prod_{i=1}^{n} (e^t p + (1-p))^{n_i} = (e^t p + (1-p))^{\sum_{i=1}^{n} n_i}$$

Entonces se sigue que:

$$Z \sim Bin(\sum_{i=1}^{n} n_i, p) \blacksquare$$

(j) Si $\forall i \; \in \{1,..,n\}$ $X_i \sim N(\mu, \sigma_i^2)$
La función generadora de momentos de una variable aleatoria $Z = \sum_{i=1}^{n} X_i$ donde cada X_i tiene una generadora de momentos $Mg_{X_i}(t) = e^{t\mu + \frac{t^2\sigma_i^2}{2}}$ está dada por:

$$Mg_Z(t) = \prod_{i=1}^{n} e^{t\mu + \frac{t^2\sigma_i^2}{2}} = e^{\sum_{i=1}^{n} t\mu + \frac{t^2\sigma_i^2}{2}} = e^{t\sum_{i=1}^{n} \mu + \frac{t^2}{2}\sum_{i=1}^{n} \sigma_i^2}$$

Entonces se sigue que

$$Z \sim N(\sum_{i=1}^{n} \mu, \sum_{i=1}^{n} \sigma_i^2) \blacksquare$$

5. Sea S_n el número de éxitos en n ensayos Bernouilli independientes. Demuestre que

$$P(|\frac{S_n}{n} - p| \geq \epsilon) \leq \frac{1}{4n\epsilon^2}$$

Digamos que $S_n = \sum_{i=1}^{n} X_i$ Veamos cual es la esperanza y la varianza de $\frac{S_n}{n}$.

$$E(\frac{S_n}{n}) = \frac{1}{n}E(S_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} p_i$$

$$Var(\frac{S_n}{n}) = \frac{1}{n^2}E(S_n) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} p_i(p_i - 1)$$

Si tenemos que los ensayos Bernoulli son identicamente distribuidos, es decir que $p_i = p$ tenemos que

$$E(\frac{S_n}{n}) = \frac{1}{n}\sum_{i=1}^{n} p = p$$

$$Var(\frac{S_n}{n}) = \frac{1}{n^2}\sum_{i=1}^{n} p(1-p) = \frac{p(1-p)}{n}$$

Entonces, por la desigualdad de Chevyshev

$$\begin{aligned} P(|\frac{S_n}{n} - E(\frac{S_n}{n})| \geq \epsilon) &\leq \frac{Var(\frac{S_n}{n})}{\epsilon^2} \\ P(|\frac{S_n}{n} - p| \geq \epsilon) &\leq \frac{\frac{p(1-p)}{n}}{\epsilon^2} \\ &= \frac{p(1-p)}{n\epsilon^2} \end{aligned}$$

Ahora hay que ver cual es el valor máximo de $p(1-p) = p - p^2$.

$$\begin{aligned} \frac{d(p-p^2)}{dp} &= 1 - 2p = 0 \\ &\implies p = \frac{1}{2} \\ \frac{d(1-2p)}{dp} &= -2 \end{aligned}$$

Por lo que $p(1-p)$ alcanza un máximo en $p = \frac{1}{2}$. Entonces, el máximo posible valor de $p(1-p)$ es

$$\frac{1}{2}(1 - \frac{1}{2}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Por lo que, para cualquier valor de p, tenemos que $p(1-p) \leq \frac{1}{4}$. Entonces, en particular

$$\frac{p(1-p)}{n\epsilon^2} \leq \frac{\frac{1}{4}}{n\epsilon^2} = \frac{1}{4n\epsilon^2}$$

Por lo que se tiene la desigualdad

$$P(|\frac{S_n}{n} - p| \geq \epsilon) \leq \frac{1}{4n\epsilon^2}$$

6. Sea S el número de águilas en 1,000,000 lanzamientos de una moneda honesta. Use (a) la desigualdad de Chebyshev y (b) el Teorema del Límite Central para estimar la probabilidad de que S esté entre 499,500 y 500,500.
   Primero veamos como se comporta S.
   X es la cantidad de águilas en un lanzamiento de esa moneda, entonces $X \sim Bernoulli(\frac{1}{2})$.
   Entonces $E(X) = \frac{1}{2}$ y $Var(X) = \frac{1}{4}$ Como S se obtiene realizando y sumando 1,000,000 veces X, entonces,

$$S = \sum_{i=1}^{1,000,000} X$$
$$E(S) = \sum_{i=1}^{1,000,000} E(X) = \sum_{i=1}^{1,000,000} \frac{1}{2} = 500,000$$
$$Var(S) = \sum_{i=1}^{1,000,000} Var(X) = \sum_{i=1}^{1,000,000} \frac{1}{4} = 250,000$$

   Pues el lanzamiento de moneda es independiente en cada ocasión.
   Ahora veamos los incisos

   (a)
$$P(499,500 \leq S \leq 500,500) = P(|S - 500,000)| \leq 500) = 1 - P(|S - 500,000)| > 500)$$

   Luego, por la desigualdad de Chebyshev, tenemos que

$$P(|S - 500,000)| > 500) < \frac{250,000}{500^2} = 1$$

   Entonces

$$P(499,500 \leq S \leq 500,500) = 1 - P(|S - 500,000)| > 500) > 1 - 1 = 0$$

   (b) Por el TLC tenem que si $X_i, \ldots, X_n$ son independientes e indenticamente distribuidas.

$$P\Big\{\frac{X_1 + \cdots + X_n - nE(X_i)}{\sqrt{nVar(X_i)}} \leq k\Big\} \to \Phi(k)$$

Cuando $n \to \infty$.
En este caso, tenemos que S es la suma de variables aleatorias independientes e indenticamente distribuidas y $n = 1,000,000$.

$$P\Big\{\frac{S - \frac{1{,}000{,}000}{2}}{\sqrt{\frac{1{,}000{,}000}{4}}} \leq k\Big\} = P\Big\{\frac{S - 500,000}{500} \leq k\Big\} = P\Big\{\frac{S}{500} - 1,000 \leq k\Big\}$$

$$= P\Big\{S \leq 500(k + 1000)\Big\} \to \Phi(k)$$

Por otra, tenemos que

$$P(499,500 \leq S \leq 500,500) = P(S \leq 500,500) - P(S \leq 499,500) + P(S = 499,500)$$

Luego,

$$\begin{aligned} 500,500 &= 500(k_1 + 1000) \implies 1,001 = k_1 + 1000 \implies k_1 = 1 \\ 499,500 &= 500(k_2 + 1000) \implies 999 = k_2 + 1000 \implies k_2 = -1 \end{aligned}$$

Por lo que

$$\begin{aligned} P(499,500 \leq S \leq 500,500) &= P(S \leq 500,500) - P(S \leq 499,500) + P(S = 499,500) \\ &= P(S \leq 500(1 + 1000)) - P(S \leq 500(-1 + 1000)) + P(S = 499,500) \\ &\approx \Phi(1) - \Phi(-1) + 0 = \Phi(1) - (1 - \Phi(1)) \\ &= 0.84134 - 1 + 0.84134 = 0.68268 \end{aligned}$$

7. Si $X \sim Gamma(n, 1)$ aproximadamente que tan grande debe ser n para que

$$P[|\frac{X}{n} - 1| > 0.01] < 0.01$$

Definamos $X_i \sim Exp(1)$, como hemos visto en clase, Gamma es una generalización de Exponencial, entonces podemos decir que:
$X = \sum_{i=1}^{n} X_i$,
Así podemos ver que:
$P[|\frac{X}{n} - 1| > 0.01] < 0.01 = P[|\frac{\sum_{i=1}^{n} X_i}{n} - 1| > 0.01] < 0.01$

$= P[|\frac{\sum_{i=1}^{n} X_i - n}{n}| > 0.01] < 0.01$

Sabemos que el Teorema del Limite Central dice que:
z = $\frac{\sum_{i=1}^{n} X_i - \mu n}{\sigma\sqrt{n}} \xrightarrow{n\to\infty} z \sim N(0,1)$

Vemos que la exponencial y varianza valen 1, es decir, $\mu = 1$ y $\sigma = 1$, así que para que nuestra expresión sea igual al teorema del límite central, necesitamos una $\sqrt{n}$ en el denominador, esto lo podemos conseguir multiplicando por $\sqrt{n}$ los dos lados de la desigualdad, esto no afecta, ya que la raíz cuadrada respeta el valor absoluto, entonces tenemos:

$= P[|\frac{\sum_{i=1}^{n} X_i - n}{\sqrt{n}}| > 0.01\sqrt{n}] < 0.01$

Vemos que es igual a la definición del Teorema del Límite Central y aplicandolo tenemos que cuando $n \to \infty$:

$P(|z| > 0.01\sqrt{n}) < 0.01$

$= (1 - P(z \leq 0.01\sqrt{n})) < 0.01$

$= (1 - \phi(0.01\sqrt{n})) < 0.01$

$= -\phi(0.01\sqrt{n}) < 0.01 - 1$

$= \phi(0.01\sqrt{n}) > 0.99$

Revisando la tabla de distribución de la normal, podemos observar que el primer valor donde ϕ es mayor a 0.99 es 2.33, entonces:

$\phi(2.33) > 0.99$ y $\phi(0.01\sqrt{n}) > 0.99$

De ahí,

$0.01\sqrt{n} = 2.33$

$\sqrt{n} = \frac{2.33}{0.01}$

$n = (\frac{2.33}{0.01})^2 = 54289$

Por lo tanto n debe ser al menos 54289 para que $P[|\frac{X}{n} - 1| > 0.01] < 0.01$

8. Sean $X_1, ..., X_{20}$ variables aleatorias independientes tales que $\forall i \ \in \{1, ..., 20\}$ $X_i \sim Poisson(\lambda)$ con media 1.

   (a) Utilice la desigualdad de Markov para obtener una cota para $P[\sum_{i=1}^{20} X_i > 15]$
   Sabemos que la desigualdad de Markov dice que:
   $P(g(x) \geq k) \leq \frac{E(g(x))}{k}$

   Ya que la esperanza de cada X_i es 1, la esperanza de la suma, es 20.
   Sustituimos en la desigualdad y tenemos que:

   $P(\sum_{i=1}^{20} X_i > 15) < \frac{E(\sum_{i=1}^{20} X_i)}{15} = \frac{20}{15}$

   Otra forma de ver el problema es:
   $P(\sum_{i=1}^{20} X_i \geq 16) \leq \frac{E(\sum_{i=1}^{20} X_i)}{16} = \frac{20}{16}$

   (b) Utilice el Teorema del Limite Central para aproximar $P[\sum_{i=1}^{20} X_i > 15]$
   Sabemos que el Teorema del Limite Central dice que:
   z $= \frac{\sum_{i=1}^{n} X_i - \mu n}{\sigma\sqrt{n}} \xrightarrow{n \to \infty} z \sim N(0, 1)$

   Ya que la esperanza de la suma es 20, al ser distribución poisson, sabemos que la varianza es igual a la esperanza, entonces la esperanza de la suma también es 20, es decir $\mu = 20$ y $\sigma = \sqrt{20}$, sustituimos en el teorema y tenemos:

z $= \frac{\sum_{i=1}^{20} X_i - 20n}{\sqrt{20n}} \xrightarrow{n\to\infty} z \sim N(0,1)$

Entonces tenemos que:
$P(z > \frac{15-20}{\sqrt{20}}) = 1 - P(z \leq \frac{15-20}{\sqrt{20}}) = 1 - \phi(\frac{15-20}{\sqrt{20}}) = 1 - \phi(-\frac{\sqrt{5}}{2}) = 1 - (1 - \phi(\frac{\sqrt{5}}{2})) = \phi(\frac{\sqrt{5}}{2})$
$\approx \phi(1.11) = 0.8665$

Por lo tanto:
$P[\sum_{i=1}^{20} X_i > 15] \approx 0.8665$

9. Sea $\{X_i\}_{i=1}^{\infty}$ una sucesión de variables aleatorias independientes. Suponga que $E(X_i) = 0$ y $Var(X_i) = \sigma_i^2$ y suponga también que:

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{\sigma_i^2}{n^2} = 0$$

Demuestre que para cualquier $\epsilon > 0$

$$P(\frac{|\sum_{i=1}^{n} X_i|}{n} > \epsilon) \to 0$$

Cuando $n \to \infty$.

Tenemos, que la ley débil de los grandes números enuncia, para una sucesión de variables aleatorias independientes $X_1, X_2, ...$, cada cual teniendo esperanza finita $\mathbb{E}(X_i) = \mu$, entonces para cualquier $\epsilon > 0$:

$$P\Big\{\Big|\frac{X_1 + X_2 + ... + X_n}{n}\Big| - \mu \geq \epsilon\Big\} \to 0$$

cuando $n \to \infty$

Luego, como por hipótesis se tiene que $\mathbb{E}(X_i) = 0$ para cada X_i, y $\mathbb{E}(X_i)$ es finita, y cada X_i tiene varianza σ_i^2, podemos aplicar la ley débil de los grandes números, de tal forma que para $X_1 + X_2 + ...$ se cumple:

$$P\Big\{\Big|\frac{X_1 + X_2 + ... + X_n}{n}\Big| - \mu \geq \epsilon\Big\} \to 0 \text{ cuando } n \to \infty$$

y como para cada X_i, $\mathbb{E}(X_i) = 0$, esto puede reescribirse como:

$$P\Big\{\Big|\frac{X_1 + X_2 + ... + X_n}{n}\Big| - 0 \geq \epsilon\Big\} \to 0 \text{ cuando } n \to \infty$$

$$P\Big\{\Big|\frac{X_1 + X_2 + ... + X_n}{n}\Big| \geq \epsilon\Big\} \to 0 \text{ cuando } n \to \infty$$

lo cual equivale a:

$$P\Big\{\frac{|X_1 + X_2 + ... + X_n|}{n} > \epsilon\Big\} \to 0 \text{ cuando } n \to \infty$$

esto puede reescribirse como:

$$P\Big\{\frac{|\sum_{i=1}^{n} X_i|}{n} > \epsilon\Big\} \to 0 \text{ cuando } n \to \infty$$

$\square$

Use lo anterior para probar que si $\{Y_i\}_{i=1}^{\infty}$ son variables aleatorias independientes tales que $Y_i \sim Bnnlli(p_i)$ entonces para cualquier $\epsilon > 0$,

$$P(|\frac{\sum_{i=1}^{n} Y_i}{n} - p(n)| \leq \epsilon) \to 1$$

Cuando $n \to \infty$ donde $p(n) = \frac{\sum_{i=1}^{n} p_i}{n}$

Tenemos que se suponen las Y_i independientes, y todas idénticamente distribuidas, como cada Y_i se distribuye Bernoulli, tenemos que $\mathbb{E}(Y_i) = p_i$ para cada Y_i.

Si denotamos a $\bar{Y} = \frac{Y_1+Y_2+...+Y_n}{n}$, entonces tenemos:

$$\mathbb{E}(\bar{Y}) = \mathbb{E}(\frac{Y_1 + Y_2 + ... + Y_n}{n}) = \frac{1}{n}\mathbb{E}(Y_1 + Y_2 + ... + Y_n)$$

y como $\mathbb{E}(Y_i) = p_i$ para cada Y_i:

$$\mathbb{E}(\bar{Y}) = \frac{\sum_{i=1}^{n} p_i}{n} = p(n)$$

así, podemos aplicar la ley débil de los grandes números a $\bar{Y}$, de tal manera que:

$$P(|\bar{Y} - \mathbb{E}(\bar{Y})| \leq \epsilon) = P\Big(\Big|\frac{\sum_{i=1}^{n} Y_i}{n} - p(n)\Big| \leq \epsilon\Big) = 1 - P\Big(\Big|\frac{\sum_{i=1}^{n} Y_i}{n} - p(n)\Big| \geq \epsilon\Big)$$

finalmente, aplicando la ley débil de los grandes números, cuando $n \to \infty$:

$$1 - P\Big(\Big|\frac{\sum_{i=1}^{n} Y_i}{n} - p(n)\Big| \geq \epsilon\Big) \to 1 - 0 = 1$$

$\square$

10. Explique por qué una variable aleatoria con distribución $Gamma(r, \lambda)$ tiene aproximadamente una distribución normal cuando r es grande.

Tenemos por resultado visto en clase, que una variable aleatoria con distribución $Gamma(r, \lambda)$ se puede ver como la suma de 'r' variables aleatorias con distribución $Exp(\lambda)$.

Cuando r es grande, puede aplicarse el teorema del límite central a la secuencia de $X_1, X_2, ...$ variables independientes e identicamente distribuidas con distribución exponencial, con lo cual, la distribución para la suma de estas variables aleatorias distribuidas $Exp(\lambda)$ (que equivalen a una v.a. distribuida $Gamma(r, \lambda)$) es aproximadamente una normal estándar $(N(0, 1))$, cuando r es grande.