Ejercicios Tema 7 - Ley de los grandes números y Teorema Central del Límite

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

Curso de Probabilidad y Variables Aleatorias con R y Python

Contenidos

1 Convergencia de variables aleatorias

1.1 Pregunta 1.

Demostrar que si la sucesión $\{X_n\}_{n=1}^{\infty}$ converge hacia X casi seguramente, la sucesión $\{X_n - X\}_{n=1}^{\infty}$ converge hacia cero casi seguramente.

1.1.1 Solución

Indicaciones para la solución:

La resolución es inmediata a partir de la definición de la convergencia casi segura.

1.2 Pregunta 2.

Demostrar que si la sucesión $\{X_n\}_{n=1}^{\infty}$ converge hacia X en probabilidad, la sucesión $\{X_n - X\}_{n=1}^{\infty}$ converge hacia cero en probabilidad.

1.2.1 Solución

Indicaciones para la solución:

La resolución es inmediata a partir de la definición de la convergencia casi segura.

1.3 Pregunta 3.

Sea $\{X_n\}_{n=1}^{\infty}$ la sucesión de variables aleatorias Bernoulli con parámetro $p_n = \frac{1}{n^2}$. Estudiar la convergencia casi segura, en probabilidad y en ley de la sucesión anterior.

1.3.1 Solución

Indicaciones para la solución:

Primero tenemos que reflexionar hacia dónde puede tender la sucesión X_n. Si $P(X_n = 1) = 1/n^2$ y $P(X_n = 0) = 1 - 1/n^2$, observamos que a medida que n aumenta, es más probable que X_n alcance el valor 0. Por tanto, “parece” que X_n tiende hacia la variable constante $X = 0$.

Para estudiar la convergencia casi segura y en probabilidad, hay que calcular la probabilidad del suceso siguiente $P(|X_n| > \epsilon)$.

El suceso $\{|X_n| > \epsilon\}$ es equivalente a decir $\{X_n = 1\}$ ya que si $X_n = 0$, $|X_n|$ no puede ser mayor que ϵ.

Por tanto:

$$P(|X_n| > \epsilon) = P(X_n = 1) = 1/n^2.$$

Para estudiar la convergencia casi segura, hay que estudiar la convergencia de la serie $\displaystyle\sum_{n=0}^{+\infty} \frac{1}{n^2}$.

Para estudiar la convergencia en probabilidad, hay que estudiar si la sucesión $\{\frac{1}{n^2}{}_{n=0}^{+\infty}$ tiende a cero o no.

Para estudiar la convergencia en ley, hay que ver si la función de distribución de X_n (ver en el tema de distribuciones notables cómo es la distribución de la variable de Bernoulli) tiende hacia la función de distribución de $X = 0$ que vale:

$F_X(x) = 0$, si $x < 0$ y $F_X(x) = 1$, si $x \geq 1$.

Pensad que si tenéis asegurada la convergencia casi segura, automáticamente tenéis la convergencia en probabilidad y en ley.

1.4 Pregunta 4.

Sea $\{X_n\}_{n=1}^{\infty}$ la sucesión de variables aleatorias con función de densidad:

$$f_{X_n}(x) = \begin{cases} \frac{nx^{n-1}}{\theta^n}, & \text{si } 0 < x < \theta, \\ 0, & \text{en caso contrario,} \end{cases}$$

donde θ es un valor real positivo. Estudiar la convergencia casi segura, en probabilidad y en ley de la sucesión anterior.

1.4.1 Solución

Indicaciones para la resolución:

Igual que en el ejercicio anterior, primero hay que reflexionar hacia dónde puede tender X_n. La función de densidad $f_{X_n}(x)$ vale 0 si x no está en $(0,\theta)$. Si $x \in (0,\theta)$ entonces $f_{X_n}(x) = (n/x)\cdot(x/\theta)^n$, como (x/θ) está entre 0 y 1, el límite de $f_{X_n}(x)$ valdrá 0 si n tiende a infinito. Ahora bien, si x tiende a θ,notemos que la función de densidad tiende a $\frac{n}{\theta}$ que tiende a infinito. Por tanto, "parece" que X_n tiende a la función constante $X = \theta$.

Entonces, estudiar la convergencia casi segura y en probabilidad es equivalente a hallar la probabilidad del suceso

$$p_n = P(|X_n - \theta| > \epsilon).$$

La probabilidad anterior puede escribirse como

$$1 - P(|X_n - \theta| \le \epsilon) = 1 - P(\theta - \epsilon < X_n < \theta + \epsilon).$$

Usad la función de densidad dada para hallar la probabilidad anterior p_n. Para estudiar la convergencia casi segura, hay que estudiar la convergencia de la serie $\sum_{n=}^{+\infty} p_n$ y para estudiar la convergencia en probabilidad hay que estudiar el límite de la sucesión p_n.

Por último, para estudiar la convergencia en ley, hay que hallar la función de distribución de X__n y ver si tiende a la función de distribución de la variable $X = \theta$.

Pensad que si tenéis asegurada la convergencia casi segura, automáticamente tenéis la convergencia en probabilidad y en ley.

1.5 Pregunta 5.

Sea $\{X_n\}_{n=1}^{\infty}$ la sucesión de variables aleatorias binomiales de parámetros n y $p_n = \frac{1}{n^2}$. Demostrar, usando la desigualdad de Chebyschev, que la sucesión $\left\{X_n - \frac{1}{n}\right\}_{n=1}^{\infty}$ converge en probabilidad hacia cero.

1.5.1 Solución

Indicaciones para la resolución:

Usando que el valor medio de las variables Xn es $E(X_n) = n \cdot p_n = \frac{1}{n}$ y que la varianza es $Var(X_n) = n \cdot p_n \cdot (1 - p_n) = \frac{1}{n} \cdot \left(1 - \frac{1}{n^2}\right)$, basta acotar la probabilidad

$$P\left(\left|X_n - \frac{1}{n}\right| > \epsilon\right) = P\left(|X_n - E(X_n)| > \epsilon\right)$$

aplicando la desigualdad de Chebyschev. Los detalles son bastante sencillos.

2 Ley de los grandes números

2.1 Pregunta 6.

Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias binomiales de parámetros m y p. 1. Demostrar que la sucesión de medias muestrales converge en probabilidad y casi seguramente hacia mp. 1. Con ayuda de `R` realizar la simulación correspondiente que comprueba la afirmación anterior. Fijar los valores siguientes $m = 10$, $p = \frac{1}{2}$ y $\epsilon = 0.05$. Considerar $k = 500$ muestras aleatorias de tamaño $N = 100$. ¿Qué efecto tienen en las simulaciones anteriores los parámetros ϵ, p y m?

2.1.1 Solución

Indicaciones para la resolución

Para resolver el primer apartado, basta aplicar la ley débil y la ley fuerte de los grandes números a la sucesión de variables X_n.

Para resolver el segundo apartado, miraos y usad los mismos pasos del script de R que tenéis en los apuntes para la simulación de las leyes de los grandes números usando una sucesión de variables de Bernoulli de parámetro $p = \frac{1}{2}$.

2.2 Pregunta 7.

Realizar el ejercicio anterior pero en lugar de suponer que las variables X_n son binomiales, suponed que son exponenciales de parámetro λ. Fijad $\lambda = 2$ en la simulación.

2.2.1 Solución

Indicaciones para la resolución

Comentarios similares al apartado anterior pero en este caso es un variaple continua exponencial.

3 Teorema Central del Límite

3.1 Pregunta 8.

El número de accidentes en un trozo de 10 Km de carretera de 2 carriles es una variable aleatoria de Poisson con media de 2 accidentes por semana. ¿Cuál es la probabilidad (aproximada) que haya menos de 100 accidentes en este trozo durante 1 año?

3.1.1 Solución

Indicaciones para la resolución:

Modelaremos el problema anterior con la variables aleatorias siguientes: sea Xi el número de accidentes en el trozo de carretera en la semana i-ésima.

La distribución de Xi será de Poisson de parámetro lambda=2. La media de Xi será 2 y la varianza también será 2.

Ahora consideremos la variable X que nos modela el número de accidentes en este trozo durante un año. Fijaos que X será

X=X1+X2+…+X52, ya que 1 año tiene 52 semanas.

Usando el Teorema Central del Límite la distribución aproximada de X será normal de media 52*media(Xi)=52*2=104 y varianza=52*Var(Xi)=104.

Nos piden P(X<100). Como sabemos la distribución aproximada de X, el cálculo de la probabilidad anterior es bastante sencillo.

3.2 Pregunta 9.

La longitud que se puede estirar un hilo de nilon se modela con una variable aleatoria exponencial con media de 1524 metros. ¿Cuál es la probabilidad (aproximada) de que la longitud media de 100 hilos esté entre 1447.8 m y 1691.64 m?

3.2.1 Solución

Es similar a los problemas anteriores.

3.3 Pregunta 10.

Las llamadas telefónicas que se reciben en un conmutador de una industria se producen como sucesos de Poisson a razón de 120 por hora. ¿Cuál es la probabilidad que lleguen entre 110 y 125 llamadas entre las 9.00 y las 10.00 de la mañana un día cualquiera?

3.3.1 Solución

Para realizar esta cuestión hemos de aplicar la aproximación de la variable de Poisson a la normal.

Sea X el número de llamadas telefónicas. X sigue la distribución de Poisson con media 120 y varianza 120.

La distribución aproximada de X será normal de media 120 y varianza 120. Nos piden P(110 <= X <=125).

Como sabemos la distribución aproximada de X el cálculo de la probabilidad anterior es bastante sencillo pero como hay que usar la continuidad de Fisher, hay que calcular considerando que X es normal la probabilidad siguiente: P(109.5 <= Z <= 125.5).

3.4 Pregunta 11.

La probabilidad de que un jugador de bàsquet enceste es p. ¿Cuántos lanzamientos tiene que hacer como mínimo (aproximadamente) para que la probabilidad de que la media de aciertos esté a distancia 0.01 de p sea de 0.99?

3.4.1 Solución

Indicaciones para la resolución:

Modelaremos el problema anterior con la variables aleatorias siguientes: Sea X_i la variable de Bernoulli de parámetro p que vale 1 si el jugador encesta y 0 en caso contrario. Si el jugador hace n lanzamientos la variable que nos da los aciertos es $X_1+. . . +X_n que sigue la distribución binomial de parámetros n y p. Nos piden hallar el n mínimo para el que se cumpla:

$$P(|\overline{X} - p| < 0.01) = 0.99,$$

donde $\overline{X} = \frac{X_1 \ldots + X_n}{n}$.

La probabilidad anterior puede escribirse como:

$$P(p - 0.01 < \overline{X} < p + 0.01) = 0.99.$$

Usando el Teorema Central del Límite, podemos decir que la distribución de $\overline{X}$ es aproximadamente normal de media p y varianza $\frac{p\cdot(1-p)}{n}$.

Por tanto, si estandarizamos la probabilidad anterior por una $Z \approx N(0, 1)$ obtenemos:

$\left(\frac{-0.01}{\sqrt{\frac{p\cdot(1-p)}{n}}}\right)$

$$P\left(\frac{-0.01}{\sqrt{\frac{p\cdot(1-p)}{n}}} < Z < \frac{0.01}{\sqrt{\frac{p\cdot(1-p)}{n}}}\right) = 0.99.$$

Usando la simetría de la distribución Z=N(0,1), podemos escribir la probabilidad anterior como:

$$2 \cdot P\left(Z < \frac{0.01}{\sqrt{\frac{p\cdot(1-p)}{n}}}\right) - 1 = 0.99$$

,

de donde:

$$P\left(Z < \frac{0.01}{\sqrt{\frac{p\cdot(1-p)}{n}}}\right) = \frac{1+0.99}{2} = 0.995.$$

De aquí deducimos, mirando el cuartil 99.5% de la normal $Z \approx N(0,1)$ que

$$\frac{0.01}{\sqrt{\frac{p\cdot(1-p)}{n}}} = 2.575.$$

Si despejamos n, obtenemos:

$$n = 2.575^2 \cdot \frac{p\cdot(1-p)}{0.01^2}$$

.

Ahora tenemos el problema que no conocemos p. Pero podemos ponernos en el "peor" de los casos. Os dejamos como ejercicio demostrar que

$p\cdot(1-p) \leq 1/4$, o sea la función $p\cdot(1-p)$ alcanza su máximo en $p=\frac{1}{2}$ para $p=[0,1]$. Para comprobarlo, basta que hagáis con R un gráfico de la función $p\cdot(1-p)$ en $(0,1)$ y veréis que en $p=\frac{1}{2}$ tiene un máximo.

Entonces el valor de n será en el "peor" de los casos:

$$n = 2.575^2 \cdot \frac{0.25}{0.01^2} = 16576.56$$

Por lo que , $n \geq 16577$.

De hecho $n \geq 16588$ si usamos muchos más decimales del cuantil 99.5% de la distribución $N(0,1)$.

3.5 Pregunta 12.

Sea $X_1,\ldots,X_n$ con $n=48$, una muestra aleatoria simple de una variable aleatoria uniforme en el intervalo $(0,a)$. Aplicando el Teorema Central del Límite, hallar la probabilidad aproximada de que $\sum\limits_{i=1}^{n} X_i > a$.

3.5.1 Solución

Indicaciones para la resolución:

La distribución aproximada de la variable suma $S = X_1+\ldots+X_n$, aplicando el Teorema Central del Límite será normal de media $n\cdot E(X_i) = 48\cdot a/2 = 24\cdot a$ y varianza $n\cdot Var(X_i) = 48\cdot\frac{a^2}{12} = 4\cdot a^2$.

Nos piden $P(S>a)$. Si estandarizamos S, tenemos que el valor de la probabilidad anterior será:

$$P(Z > (a-24\cdot a)/(\sqrt{4}\cdot a) = P(Z > -23\cdot\frac{a}{2\cdot a}) = P(Z > -23/2)$$

que aproximadamente vale 1.