Suma y promedio de variables aleatorias independientes. Ley de los grandes números. Teorema central del Limite.

Teorema (Desigualdad de Tchebycheff).
Sea X una v.a. con cualquier distribución, con esperanza y varianza finitas. Sea a>0. Entonces:

$$\mathbf{P(\ |X - E(X)| \geq a\ )\ \leq\ \frac{Var(X)}{a^2}}$$

Comentario: observar que la desigualdad de Tchebycheff es equivalente a:

$$P(\ |X - E(X)| < a\ )\ \geq\ 1 - \frac{Var(X)}{a^2}$$

Ejemplo de aplicación: Sea X una v.a. con cualquier distribución con esperanza y varianza finitas. Llamemos μ a su esperanza y σ^2 a su varianza. Dar una cota para

a) $P(\mu-2\sigma \leq X \leq \mu+2\sigma)$

Si sabemos que X~Normal, cuánto vale esta probabilidad?

b) idem a) para $P(\mu-3\sigma \leq X \leq \mu+3\sigma)$

c) idem a) para $P(\mu-\sigma \leq X \leq \mu+\sigma)$

Esperanza y varianza de sumas de vs. as. independientes

Dijimos que si $X_1,...,X_n$ son vs. as.

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

Si $X_1,...,X_n$ son vs. as. independientes

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

Proposición : Sean $X_1,...,X_n$ son vs. as. independientes, todas con la misma esperanza y la misma varianza finita. Llamemos μ a $E(X_i)$ y σ^2 a la $Var(X_i)$. Sea

$$S = \sum_{i=1}^{n} X_i \qquad \text{y} \qquad \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

entonces

$$E(S) = n\mu \qquad , \qquad Var(S) = n\sigma^2$$

$$E(\overline{X}) = \mu \qquad , \qquad Var(\overline{X}) = \frac{\sigma^2}{n}$$

Nota: obsérvese que el caso $X_1,...,X_n$ vs. as. independientes igualmente distribuidas (i.i.d.) con esperanza y varianza finita cumple las condiciones de la proposición.

Convergencia de una sucesión de variables aleatorias.

Una definición (no es la única que se usa) es la siguiente:

Definición de convergencia en probabilidad.

Sea $Y_1,...,Y_n,...$ una sucesión de variables aleatorias. Se dice que

$$Y_n \xrightarrow[n \to \infty]{} Y \quad \textbf{en probabilidad}$$

si para cualquier $\varepsilon>0$ se cumple que

$$\lim_{n \to \infty} P(|Y_n - Y| > \varepsilon) = 0$$

Teorema (Ley de los Grandes Números, versión de Tchebycheff**)**

Sean $X_1,...,X_n$ son vs. as. independientes, todas con la misma esperanza y la misma varianza finita. Llamemos μ a $E(X_i)$ y σ^2 a la $Var(X_i)$. Sea

$$\overline{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$$

Entonces

$$\overline{X}_n \xrightarrow[n\to\infty]{} \mu \quad \text{en probabilidad}$$

La demostracion del teorema anterior sale facil de la desigualdad de Tchebycheff. Luego se probó que no es necesario suponer varianza finita, como lo muestra el teorema siguiente.

Teorema (Ley de los Grandes Números, versión de Khintchine, 1929**)**

Sean $X_1,...,X_n$ son vs. as. i.i.d., todas con esperanza finita. Llamemos μ a $E(X_i)$

Entonces

$$\overline{X}_n \xrightarrow[n\to\infty]{} \mu \quad \text{en probabilidad}$$

Distribución de la suma de variables independientes.

Proposición :

Sean X e Y variables aleatorias independientes.

a) Si $X \sim Bin(n_1,p)$, $Y \sim Bin(n_2,p)$ entonces $X+Y \sim Bin(n_1+n_2,p)$.

b) Si $X \sim Poisson(\lambda_1)$, $Y \sim Poisson(\lambda_2)$, entonces $X+Y \sim Poisson(\lambda_1+\lambda_2)$

c) Si $X \sim N(\mu_1,\sigma_1^2)$, $Y \sim N(\mu_2,\sigma_2^2)$ entonces $X+Y \sim N(?,?)$ (rellenar los ?)

d) Si $X \sim \Gamma(\alpha_1, \lambda)$, $Y \sim \Gamma(\alpha_2, \lambda)$ entonces $X+Y \sim \Gamma(\alpha_1+\alpha_2, \lambda)$

El inciso a) es intuitivamente razonable, verdad? Los otros no son intuitivos.

Comentario: Esta propiedad de que la distribución de la suma de 2 vs as independiente "pertenece a la misma familia" no es general. Por ejemplo, si X e Y son vs. as. independientes, cada una con distribución uniforme en el intervalo [a,b], X+Y no tiene distribución uniforme.

Proposición (generalización de la proposición anterior)

Sean $X_1,...,X_m$ variables aleatorias independientes.

a) Si $X_i \sim Bin(n_i,p)$ entonces $\sum_{i=1}^{m} X_i \sim Bin(\sum_{i=1}^{m} n_i, p)$.

b) Si $X_i \sim Poisson(\lambda_i)$ entonces $\sum_{i=1}^{m} X_i \sim Poisson(\sum_{i=1}^{m} \lambda_i)$.

c) Si $X_i \sim N(\mu_i,\sigma_i^2)$, entonces $\sum_{i=1}^{m} X_i \sim N(\sum_{i=1}^{m} \mu_i, \sum_{i=1}^{m} \sigma_i^2)$.

d) Si $X_i \sim \Gamma(\alpha_i, \lambda)$ entonces $\sum_{i=1}^{m} X_i \sim \Gamma(\sum_{i=1}^{m} \alpha_i, \lambda)$

Vamos a enunciar dos resultados sobre la districión de la suma y el promedio de variables i.i.d.

Teorema 1:

Sean $X_1,...,X_n$ son vs. as. i.i.d. $N(\mu,\sigma^2)$. Sea

$$S = \sum_{i=1}^{n} X_i \quad \text{y} \quad \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

entonces

$$S \sim N(n\mu, n\sigma^2), \quad \overline{X} \sim N(\mu, \sigma^2/n)$$

Este teorema es consecuencia del inciso b) de la proposición anterior.

Teorema 2 (con la tesis escrita en forma no rigurosa)

Sean $X_1,...,X_n$ son vs. as. i.i.d. **con cualquier distribución**, con esperanza y varianza finitas. Llamemos μ a la $E(X_i)$ y $\sigma^2=Var(X_i)$. Sea

$$S_n = \sum_{i=1}^{n} X_i \quad \text{y} \quad \overline{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$$

entonces

$$S_n \underset{(a)}{\sim} N(\;,\;) \qquad \overline{X}_n \underset{(a)}{\sim} N(\;,\;)$$

cuando n es grande.

Usamos $\underset{(a)}{\sim}$ como abreviatura de “tiene distribución aproximadamente”

Comentarios:

- Cuál es más llamativo el teorema 1 o el teorema 2?
- El Teorema 2 se llama **Teorema Central del Límite**.

La tesis del TCL se escribe formalmente así:

Sea $Z_n = \dfrac{\overline{X}_n - \mu}{\sqrt{\sigma^2 / n}}$, o lo que es lo mismo, sea $Z_n = \dfrac{S_n - n\mu}{\sqrt{n\sigma^2}}$

entonces

$$\lim_{n \to \infty} F_{Z_n}(z) = \Phi(z)$$

“Receta”: En la mayoría de las aplicaciones, la aproximación que dá el TCL es satisfactoria para $n \geq 30$.

- Como se obtuvo esta receta?: empiricamente.

- Por qué la distribución normal es muy usada? No tanto porque sirva de modelo para algunas variables aleatorias, sino gracias al TCL. Aunque una variable aleatoria tenga una distribución que no sea gaussiana, la suma o el promedio de muchas variables tiene distribución aproximadamente gaussiana. Ademas hay versiones del TCL con menos hipotesis que la version que enunciamos.

Teorema (aproximacion de la distibución binomial por la normal)

(con la tesis escrita en forma no rigurosa). Sea X~ Bin(n,p). Si n es “grande”

$$\mathbf{X \underset{(a)}{\sim} N(np,\, np(1-p))}$$

Comentarios:

- La aproximación normal a la binomial es satisfactoria cuando $np \geq 5$ y $n(1-p) \geq 5$.
- Para valores de n que cumplen la condición anterior, pero no son “muy grandes” la “corrección por continuidad” mejora mucho la aproximación.