Mitjana mostral i llei dels grans nombres

Consideram una v.a. X associada a un experiment aleatori.
(Per exemple, X=nombre de cares en el llançament d'una moneda, llavors $\Omega_X = \{0, 1\}$).

Repetim n vegades l'experiment i denotam X_i la v.a. associada a la repetició i-èssima de l'experiment. (Observem que les X_i estan idènticament distribuïdes ja que totes estan definides de manera idèntica per al mateix experiment. A més, si les repeticions de l'experiment són independents tenim que les X_i són i.i.d, i per tant $E(X_i) = E(X)$ i $\mathrm{Var}(X_i) = \mathrm{Var}(X)$).

Definim la **mitjana mostral de** X com:

$$M_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

M_n és una nova v.a. que té les següents **propietats**:

- $E(M_n) = E(X)$
- $\mathrm{Var}(M_n) = \dfrac{\mathrm{Var}(X)}{n}$

Interpretació: Les anteriors propietats ens diuen que la mitjana dels valors de M_n és igual a la mitjana de X i que la dispersió dels valors de M_n tendeix cap a zero quan n augmenta.

La relació entre la mitjana (esperança) i la dispersió (variància) dels valors d'una variable aleatòria qualsevol vé donada per la **desigualtat de Txebitxeff**:

$$P(|Y - E(Y)| \geq \varepsilon) \leq \frac{\mathrm{Var}(Y)}{\varepsilon^2} \qquad \forall \varepsilon > 0 \qquad \forall \text{v.a. } Y$$

que també se pot escriure com:

$$P(|Y - E(Y)| < \varepsilon) \geq 1 - \frac{\mathrm{Var}(Y)}{\varepsilon^2} \qquad \forall \varepsilon > 0 \qquad \forall \text{v.a. } Y$$

Aplicant l'anterió relació al cas de la mitjana mostral tenim:

$$P(|M_n - E(M_n)| < \varepsilon) \geq 1 - \frac{\mathrm{Var}(M_n)}{\varepsilon^2} \quad \Longleftrightarrow \quad P(|M_n - E(X)| < \varepsilon) \geq 1 - \frac{\mathrm{Var}(X)}{n\varepsilon^2} \qquad \forall \varepsilon > 0$$

Com a conseqüència d'aquesta relació, tenim que si n és molt gran ($n \to \infty$):

$$\lim_{n\to\infty} P(|M_n - E(X)| < \varepsilon) = 1 \qquad \forall \varepsilon > 0$$

Aquesta expressió es coneix com **Llei dèbil dels grans nombres** i ens diu que és segur (amb probabilitat 1) que si repetim un nombre molt gran de vegades un experiment, la mitjana mostral serà igual a la mitjana teòrica de la v.a. X.

Exemple 8:

(Exercici 10). Suposem que el 10% dels votants estan a favor d'una certa legislació. Es fa una enquesta entre la població i s'obté una freqüència relativa $f_n(A)$ com una estimació de la proporció anterior. Determinau, aplicant la desigualtat de Txebicheff, quants de votants s'haurien d'enquestar perquè la probabilitat que $f_n(A)$ difereixi de 0.1 menys de 0.02 sigui al menys 0.95 Què podem dir si no coneixem el valor de la proporció?

Exemple 9:

(Exercici 11). Es llança a l'aire un dau regular 100 vegades. Aplicau la desigualtat de Txebicheff per obtenir una fita de la probabilitat que el nombre total de punts obtinguts estigui entre 300 i 400.

Exemple 10:

(Exercici 12). Es sap que, en una població, la talla dels individus mascles adults és una variable aleatòria X amb mitjana $\mu_x = 170$ cm i desviació típica $\sigma_x = 7$ cm. Es tria una mostra aleatòria de 140 individus. Calculau la probabilitat que la mitjana mostral $\overline{x}$ difereixi de μ_x en menys d'1 cm.

Problemes proposats: 9, 13, 14