Indicații - Teoreme limită

• Inegalitatea lui Cebâșev se aplică în probleme în care se cer limita (marginea) superioară sau inferioară a unei probabilități, estimarea unei probabilități sau eventual să se arate că o anumită probabilitate este mai mare sau mai mică decât un nr. dat.

• Teorema lui Cebâșev

Media aritmetică a unui număr suficient de mare de variabile aleatoare independente ia, cu o probabilitate mare, valori apropiate de un număr constant, $\frac{1}{n}\sum_{i=1}^{n} E(X_i)$.

Consecință: Dacă $(X_n)_n$ este un șir de variabile aleatoare care sunt independente două câte două, ale căror dispersii sunt mărginite de o aceeași constantă și au aceeași medie m, atunci șirul $\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)_n$ converge în probabilitate către m.

Dem: $\frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} m = m$

Aplicând th. lui Cebâșev, obținem:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m\right| < \varepsilon\right) = 1, \text{ adică}$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} m \ (n\to\infty)$$

Observație: Consecința justifică regula mediei aritmetice aplicată în teoria măsurătorilor, care constă în măsurarea repetată în aceleași condiții, de n ori a mărimii fizice m, obținând n rezultate $X_1, X_2, \ldots X_n$. Drept valoare aproximativă pt. m se ia media aritmetică a rezultatelor, adică $m \simeq \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$

Dacă măsurătorile sunt lipsite de erori sistematice (adică $E(X_i) = m$, $i=\overline{1,n}$), atunci, conform legii numerelor mari pt. n suficient de mare cu o probabilitate oricât de aproape de valoarea 1, putem obține o valoare oricât de apropiată de valoarea m.

Teorema lui Bernoulli

Putem spune că este o consecință a teoremei lui Cebîșev. Ea explică stabilitatea frecvenței relative pt. un număr mare de experimente.

Aplicarea teoremei integrale a lui Moivre-Laplace

(1.) Se fac n experimente independente, în fiecare experiment probabilitatea evenimentului A fiind p. Care este probabilitatea ca numărul de apariții ν_n ale even. A să fie cuprins între K_1 și K_2, $0 \le K_1 < K_2 \le n$?

Numărul de apariții ν_n ale even. A are repartiția binomială $Bi(n,p)$.

Din teorema integrală a lui Moivre-Laplace avem:

$$P(K_1 \le \nu_n \le K_2) = P\left(\frac{K_1 - np}{\sqrt{npq}} \le \frac{\nu_n - np}{\sqrt{npq}} \le \frac{K_2 - np}{\sqrt{npq}}\right) =$$

$$= \Phi\left(\frac{K_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{K_1 - np}{\sqrt{npq}}\right), \text{ deoarece var. } \frac{\nu_n - np}{\sqrt{npq}} \sim N(0,1)$$

(2.) Sunt date numerele n, p și ε. Să se determine probabilitatea ca frecvența relativă de apariție a even. A să nu se abată în modul de la probabilitatea p a evenimentului A cu mai mult decât ε.

$$P\left(\left|\frac{\nu_n}{n} - p\right| \le \varepsilon\right) = P\left(-\varepsilon \le \frac{\nu_n}{n} - p \le \varepsilon\right) = P\left(-\varepsilon \le \frac{\nu_n - np}{n} \le \varepsilon\right)$$

$$= P\left(-\varepsilon \cdot \sqrt{\frac{n}{pq}} \le \underbrace{\frac{\nu_n - np}{\sqrt{npq}}}_{\sim N(0,1)} \le \varepsilon \cdot \sqrt{\frac{n}{pq}}\right) = \Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) - \Phi\left(-\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) =$$

$$= 2\,\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) - 1$$

3. Sunt date numerele p, ε și α. Să se determine numărul minim n de experimente care să asigure cu probabilitatea nu mai mică decât α ca frecvența relativă de apariție a even. A să nu se abată în modul de la probabilitatea p de apariție a even. A cu mai mult decât ε.

Dacă dorim o abatere destul de mică, ne așteptăm ca numărul experimentelor să fie destul de mare. Așadar, căutăm n pentru care

$$P\left(\left|\frac{\nu_n}{n}-p\right|\leq\varepsilon\right)\geq\alpha$$

Ținând cont de problema 2, numărul căutat rezultă din

$$2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right)-1\geq\alpha$$

4. Cu probabilitatea α și numărul de experimente n să se determine intervalul în care se va situa frecvența relativă de apariție a unui even. A de probabilitate p.

Căutăm un ε pentru care $P\left(\left|\frac{\nu_n}{n}-p\right|<\varepsilon\right)=\alpha$, deci din pb. 2 avem $P\left(\left|\frac{\nu_n}{n}-p\right|<\varepsilon\right)=2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right)-1=\alpha$

Numărul ε rezultă din ecuația $\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right)=\frac{1+\alpha}{2}$

Deci intervalul în care se va situa frecvența este $(p-\varepsilon, p+\varepsilon)$.

Inegalitatea lui Cebîşev. Exemplu:

Dacă în $P(|X-E(X)|<\varepsilon) \geq 1-\frac{D^2(X)}{\varepsilon^2}$, înlocuim ε cu $t\sigma$,

atunci $P(|X-E(X)|<t\sigma) \geq 1-\frac{\sigma^2}{t^2\sigma^2} = 1-\frac{1}{t^2}$

Luând $t=3$ avem

$$P(|X-E(X)|<3\sigma) \geq \frac{8}{9}$$

Constatăm că valorile var. aleat. X, drept rezultat al experimentului, nimeresc în intervalul $(E(X)-3\sigma, E(X)+3\sigma)$ cu probabilități cuprinse între $\frac{8}{9}$ şi 1 (vezi şi anexa).

Scanned with CamScanner

Anexă: Repartiția normală

Fie $X \sim N(m, \sigma)$. Atunci $Y = \frac{X-m}{\sigma}$ se numește <u>variabilă aleatoare normată</u>.

$$E(Y) = E\left(\frac{X-m}{\sigma}\right) = \frac{1}{\sigma} E(X-m) = \frac{1}{\sigma}\left(\underbrace{E(X)}_{= m} - m\right) = 0$$

$$D^2(Y) = D^2\left(\frac{X-m}{\sigma}\right) = \frac{1}{\sigma^2} D^2(X-m) = \frac{1}{\sigma^2} \underbrace{D^2(X)}_{= \sigma^2} = 1$$

Dacă var. aleat. $X \sim N(m, \sigma)$, atunci var. aleat. $Y = \frac{X-m}{\sigma} \sim N(0,1)$.

Def: Valoarea x_p a unei var. aleat. continue X pentru care $F(x_p) = P(X < x_p) = p$ se numește <u>cuantila de ordinul p</u>.

Fie x_p cuantila de ordinul p în cazul repartiției $N(0,1)$, adică $\Phi(x_p) = p$ (unde Φ este fct. lui Laplace)

Atunci $\Phi(-x_p) = 1 - \Phi(x_p) = 1 - p$, dar $\Phi(x_{1-p}) = 1 - p$.

Deci $x_{1-p} = -x_p$

Dacă $X \sim N(m, \sigma)$ și $x_1, x_2 \in \mathbb{R}$, atunci

$$P(x_1 \le X \le x_2) = \Phi\left(\frac{x_2 - m}{\sigma}\right) - \Phi\left(\frac{x_1 - m}{\sigma}\right)$$

Dem: $P(x_1 \le X \le x_2) = P\left(\frac{x_1 - m}{\sigma} \le \frac{X-m}{\sigma} \le \frac{x_2 - m}{\sigma}\right) =$

not $= Y \sim N(0,1)$

$= P\left(\frac{x_1 - m}{\sigma} \le Y \le \frac{x_2 - m}{\sigma}\right) = \Phi\left(\frac{x_2 - m}{\sigma}\right) - \Phi\left(\frac{x_1 - m}{\sigma}\right)$.

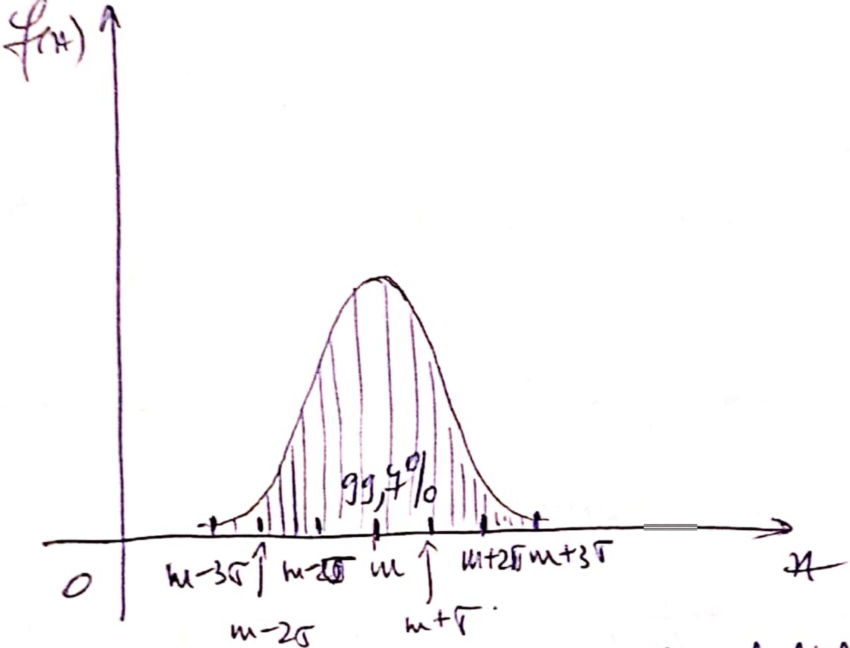
În particular, $P(|X - m| < k\sigma) = 2\Phi(k) - 1$

Dem: $P(|X - m| < k\sigma) = P(m - k\sigma < X < m + k\sigma) =$
$= P\left(-k < \underbrace{\frac{X-m}{\sigma}}_{\sim N(0,1)} < k\right) = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1$

Dacă luăm $K = 3$, obținem

$$P(|X-m|<3\sigma) = 2\Phi(3) - 1 = 2\cdot 0,9987 - 1 = 0,9974 \quad (*)$$

Egalitatea (*) exprimă faptul că practic aproape toate valorile var. X cad în intervalul $(m-3\sigma, m+3\sigma)$, așa-numita „regulă a celor trei sigma".



Exemplu: Să considerăm un ansamblu statistic (populație) de rezistențe ale căror valori sunt repartizate după o lege normală $N(200\Omega, 8\Omega)$. Luând la întâmplare 100 dintre aceste rezistențe, le controlăm una câte una. Care este probabilitatea de a se abate cu mai mult de 10Ω de la valoarea nominală de 200Ω?

Soluție: Notăm cu X var. aleat. considerată.

Aplicăm formula $P(|X-m|<K\sigma) = 2\Phi(K) - 1$, unde

$m=200$, $\sigma=8$, $K\sigma=10 \Rightarrow K=1,25$

$$P(|X-200|<10) = 2\Phi(1,25) - 1 = 2\cdot 0,8944 - 1 = 0,7888$$

iar $q = 1 - p = 1 - 0,7888 = 0,2112 = 21,12\%$.