Tema 4

Sumes de variables aleatòries

4.1 Distribució de la mitjana mostral

Sigui X una variable aleatoria amb $E(X) = \mu$ i $\mathrm{Var}(X) = \sigma^2 < \infty$. Siguin $X_1, \ldots, X_n, \ldots$ variables aleatorias independents totes amb la mateixa distribució que X. Definim la variable aleatoria MITJANA MOSTRAL de $X_1, \ldots, X_n, \ldots$ com la seva mitjana aritmètica:

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

Proposició 4.1.1.

$$E(\bar{X}_n) = \mu, \ \ Var(\bar{X}_n) = \frac{\sigma^2}{n}.$$

Demostració:

$$E(\bar{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{n\mu}{n} = \mu.$$

De la mateixa forma, i tenint en compte la independència,

$$\mathrm{Var}(\bar{X}_n) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} E(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

□

Exemple 4.1.1.

En el cas de tenir n variables aleatorias independents $N(\mu, \sigma^2)$, la seva suma és també normal (ho vam veure al tema 3). Per tant, la seva mitjana també ho és:

$$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Leftrightarrow \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0,1).$$

$\square$

Exemple 4.1.2.
En el cas de tenir n variables aleatorias independents de Bernoulli de paràmetre p, la seva suma $S_n = \sum_{i=1}^n X_i \sim B(n,p)$ i la mitjana mostral és

$$\bar{X}_n = \frac{S_n}{n},$$

mesura la proporció mostral d'èxits. Per això també es fa servir la notació

$$\hat{p}_n = \bar{X}_n = \frac{S_n}{n}.$$

Ara acabem de veure que

$$E(\hat{p}_n) = p,\ \mathrm{Var}(\hat{p}_n) = \frac{p(1-p)}{n}.$$

Això coincideix amb quelcom que ja sabíem. Recordeu que $E(S_n) = np$ i $\mathrm{Var}(S_n) = np(1-p)$. Per tant

$$E(\hat{p}_n) = E\left(\frac{S_n}{n}\right) = \frac{np}{p} = p,\ \mathrm{Var}(\hat{p}_n) = \mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

$\square$

4.2 Desigualtats: Markov, Tchebychev, Jensen, Chernoff

Les desigualtats de Tchebychev i de Markov proporcionen unes fites superiors per a la probabilitat de la cua d'una distribució en el cas centrat i en el cas no centrat, respectivament.

Teorema 4.2.1 (Desigualtat de Markov)**.** *Donada una variable aleatòria positiva X (és a dir: $P(X \geq 0) = 1$) amb $E(X) < \infty$, llavors per tot $t > 0$ es té*

$$P(X \geq t) \leq \frac{E(X)}{t}.$$

Demostració:

1. Cas discret:

$$E(X) = \sum_{x \in \mathcal{X}} x\, P(X = x) = \underbrace{\sum_{x<t} x\, P(X = x)}_{\geq 0} + \sum_{x \geq t} x\, P(X = x)$$

$$\geq \sum_{x \geq t} x\, P(X = x) \geq t \sum_{x \geq t} P(X = x) = t\, P(X \geq t).$$

2. Cas continu:

$$E(X) = \int_0^\infty x\, f(x)\, dx = \int_0^t x\, f(x)\, dx + \int_t^\infty x\, f(x)\, dx$$

$$\geq \int_t^\infty x\, f(x)\, dx \geq t \int_t^\infty f(x)\, dx = t\, P(X \geq t).$$

$\square$

Observacions:

- Aquesta desigualtat té interès quan t és gran, al menys $t \geq E(X)$. Fixem-nos que si $t \leq E(X)$ tindríem $\frac{E(X)}{t} \geq 1$, però d'altra banda sempre es compleix que $P(X \geq t) \leq 1$. Per tant, en aquest cas la fita no aporta informació nova.

- És important adonar-se que aquesta desigualtat és independent de la llei que segueix la variable aleatòria X.

- Si fem $t = kE(X)$ tenim que

$$P(X \geq kE(X)) \leq \frac{1}{k}, \text{ per tot } k > 0.$$

Exemple 4.2.1.
Sigui X una variable aleatòria positiva i contínua amb $E(X) = 1.2$. Ens preguntem quin valor hauria de prendre t per tal que puguem assegurar que $P(X \geq t) \leq 0.01$.

Com que $P(X \geq t) \leq \frac{E(X)}{t} = \frac{1.2}{t}$, farem que $\frac{1.2}{t} \leq 0.01$, i d'aquí: $t \geq 120$.

Per tant a la cua de la distribució a la dreta del valor $t = 120$ li correspon una probabilitat inferior a l'1%.

$\square$

Exemple 4.2.2.
Sigui X variable aleatòria tal que $X \sim \text{Poisson}(1)$.

D'una banda tenim que $P(X \geq t) \leq \frac{E(X)}{t} = \frac{1}{t}$.

D'altra banda $P(X \geq t) = P(X^2 \geq t^2) \leq \frac{E(X^2)}{t^2} = \frac{2}{t^2}$, on hem usat que X és positiva en la primera igualtat.

Llavors, per exemple:

$$P(X \geq 3) \leq \frac{1}{3} = 0.333,$$

$$P(X \geq 3) \leq \frac{2}{9} = 0.222.$$

Però en realitat $P(X \geq 3) = 1 - P(X \leq 2) = 0.197$.

$\square$

Teorema 4.2.2 (Desigualtat de Tchebychev)**.** *Sigui X una variable aleatòria amb $Var(X) < \infty$. Llavors, per tot $t > 0$,*

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

Demostració: Definim $Y = (X - E(X))^2$ variable aleatòria positiva amb $E(Y) = Var(X)$. Aleshores, per la desigualtat de Markov, tenim:

$$P(|X - E(X)| \geq t) = P((X - E(X))^2 \geq t^2)$$
$$= P(Y \geq t^2) \leq \frac{E(Y)}{t^2} = \frac{Var(X)}{t^2}.$$

$\square$

Una altra forma d'expressar la desigualtat és

$$P\left(\frac{(X - E(X))^2}{\sigma^2} \geq t^2\right) \leq \frac{1}{t^2},$$

que ens dóna una fita universal sobre la desviació de $|X-\mu|$ en termes de σ:

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Exemple 4.2.3.

Volem trobar una fita superior per a la probabilitat que una variable aleatòria difereixi de la seva mitjana més de 3 desviacions estàndard.

Usem la notació següent: $\sigma = \sqrt{Var(X)}$ i $\mu = E(X)$.

Llavors: $P(|X-\mu| \geq 3\sigma) \leq \frac{Var(X)}{9\sigma^2} = \frac{\sigma^2}{9\sigma^2} = \frac{1}{9} = 0.11$.

Això ens diu que, per a qualsevol variable aleatòria amb μ i σ finites, l'interval $(\mu - 3\sigma, \mu + 3\sigma)$ conté almenys $\frac{8}{9}$ de la probabilitat. És a dir:

$$P(-3\sigma < X-\mu < 3\sigma) = P\left(-3 < \frac{X-\mu}{\sigma} < 3\right) \geq \frac{8}{9}.$$

La probabilitat exacta depèn de la distribució de X. $\square$

Exemple 4.2.4.

Quina hauria d'ésser la grandària d'una mostra d'una població, amb mitjana μ desconeguda i variància 4, si volem que la distància de la mitjana mostral a la mitjana poblacional fos inferior a 1 amb probabilitat superior a 0.99?

Tenim $X_1, \ldots, X_n$ amb $E(X_i) = \mu$ i $Var(X_i) = 4$ per $i = 1, \ldots, n$.

La mitjana poblacional és $\overline{X}_n$, $E(\overline{X}_n) = \mu$ i $Var(\overline{X}_n) = \frac{4}{n}$.

Volem calcular

$$P(|\overline{X}_n - \mu| < 1) \geq 0.99,$$

o equivalentment

$$P(|\overline{X}_n - \mu| \geq 1) \leq 0.01.$$

Però $P(|\overline{X}_n - \mu| \geq 1) \leq \frac{Var(\overline{X}_n)}{1} = \frac{\sigma^2}{n} = \frac{4}{n}$.

Llavors fent $\frac{4}{n} \leq 0.01$ obtenim $n \geq 400$. $\square$

Desigualtat de Jensen

La desigualtat de Jensen relaciona l'esperança de transformacions convexes d'una variable aleatòria amb la mateixa transformació feta amb l'esperança. Recordem la definició de funció convexa.

Definició 4.2.3. *Una funció $g : A \subseteq \mathbb{R} \to \mathbb{R}$ és convexa si per a tots $a, b \in A$, $a < b$, i per a tot $\alpha \in [0, 1]$, es té que*

$$g(\alpha a + (1 - \alpha)b) \leq \alpha g(a) + (1 - \alpha)g(b).$$

Proposició 4.2.4. *Una funció $g : A \subseteq \mathbb{R} \to \mathbb{R}$ és convexa si i només si per a tot $a \in A$, existeix una recta que pasa pel punt $(a, g(a))$ i que deixa la gràfica de la funció g per sobra seu. És a dir, existeix una recta*

$$\{(x, g(a) + \lambda(x - a)),\ x \in \mathcal{A}\}$$

tal que

$$g(x) \geq g(a) + \lambda(x - a) per\ a\ tot\ x \in \mathcal{A}.$$

Teorema 4.2.5 (Desigualtat de Jensen)**.** *: Sigui X v.a. amb suport $A \subseteq \mathbb{R}$ i $E(|X|) < \infty$. Sigui $g : a \subseteq \mathbb{R} \to \mathbb{R}$ una funció convexa, amb $E(|g(X)|) < \infty$. Aleshores*

$$E(g(X)) \geq g(E(X)).$$

Demostració: Com que g és convexa, existeix λ tal que $g(x) \geq g(a) + \lambda(x - a)$ per a tots $a, x \in \mathcal{A}$. En particular aquesta desigualtat es verifica si $x = X(\omega)$ per a tot $\omega \in \Omega$ i si prenem $a = E(X)$. Per tant podem escriure que

$$g(X) \geq g(E(X)) + \lambda(X - E(X)).$$

Prenent esperances es mantenen les desigualtas:

$$E(g(X)) \geq E[g(E(X))+\lambda(X-E(X))] = g(E(X))+E(X-E(X)) = g(E(X)),$$

com volíem provar. □

Exemple 4.2.5.

Sigui X variable aleatòria amb $E(X^2) < \infty$ (aleshores $E(X) < \infty$) i sigui $g(x) = x^2$, que és una funció convexa. La desigualtat de Jensen diu que

$$E(X^2) \geq E(X)^2 \Rightarrow V(X) = E(X^2) - E(X)^2 \geq 0.$$

Tenim així una altra prova de que la variància d'una variable aleatòria és sempre no negativa.

□

Desigualtat de Chernoff

4.3 Convergència de successions de variables aleatorias

Hem vist dues variables aleatorias, $\bar{X}_n$ i $\hat{p}_n$, les distribucions de les quals depenen del nombre n d'observacions disponibles. Ens plantegem ara quin és el comportament límit d'aquestes variables aleatorias (o, millor dit, d'aquestes successions de variables aleatorias) quan el nombre d'observacions n tendeix cap a infinit. Més en general, siguin X una variable aleatoria i $\{X_n\}_n$ una successió de variables aleatorias. Volem estudiar quan i en quin sentit podem dir que

$$X_n \to X \text{ quan } n \to \infty.$$

Observeu que tant X_n, $n \geq 1$, com X són funcions de l'espai mostral Ω en $\mathbb{R}$. Per tant els conceptes de convergències de variables aleatorias tindran a veure amb els tipus de convergències de successions de funcions (convergència puntual, en norma, en mesura, ...).

4.3.1 Convergència en distribució

Definició 4.3.1. *Siguin $X, X_1, X_2, ...$ variables aleatorias amb funcions de distribució $F_X = \Pr(X \leq x)$, $F_n = \Pr(X_n \leq x)$, $n \geq 1$. Direm que $X_n \to X$ quan $n \to \infty$* EN DISTRIBUCIÓ *(o* EN LLEI*, o* DÈBILMENT*) si*

$$F_n(x) \to F_X(x) \text{ quan } n \to \infty$$

per a tot x on la funció de distribució de X, $F_X(x)$, sigui contínua, i escriurem

$$X_n \xrightarrow{D} X.$$

Exemple 4.3.1.

Sigui $X_n \sim N(\mu_n = \frac{1}{n}, 1)$, $n \geq 1$. Aleshores

$$X_n \xrightarrow{D} Z \sim N(0,1).$$

En efecte,

$$F_n(x) = \Pr(X_n \leq x) = \Pr\left(X_n - \frac{1}{n} \leq x - \frac{1}{n}\right) \stackrel{X_n-(1/n)\sim N(0,1)}{=}$$

$$\Pr\left(Z \leq x - \frac{1}{n}\right) = F_Z\left(Z \leq x - \frac{1}{n}\right) \underset{n\to\infty}{\longrightarrow} F_Z(x),$$

per a tot $x \in \mathbb{R}$. Observeu que el conjunt de punts de continuïtat de F_Z és tota la recta real.

□

Exemple 4.3.2.

Sigui $X_n \sim \text{Geom}(p_n = \frac{\lambda}{n})$, $\lambda > 0$, per a $n \geq \lambda$. Aleshores

$$\frac{X_n}{n} \xrightarrow{D} X \sim \text{Exp}(\lambda).$$

En efecte, per a qualsevol $t \geq 0$,

$$F_{X_n/n}(t) = \Pr\left(\frac{X_n}{n} \leq t\right) = \Pr(X_n \leq nt) = \Pr(X_n \leq \lfloor nt \rfloor) = 1 - \Pr(X_n > \lfloor nt \rfloor)$$

$$= 1 - \left(1 - \frac{\lambda}{n}\right)^{\lfloor nt \rfloor} = 1 - \left[\left(1 - \frac{\lambda}{n}\right)^n\right]^{\frac{\lfloor nt \rfloor}{n}} \underset{n\to\infty}{\longrightarrow} 1 - e^{-\lambda t} = F_{\text{Exp}(\lambda)}(t).$$

A l'hora de prendre el límit quan n tendeix cap a infinit, hem aplicat els següents resultats:

$$\lim_{n\to\infty}\left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda},$$

$$\lim_{n\to\infty}\frac{\lfloor nt \rfloor}{n} = t \lim_{n\to\infty}\frac{\lfloor nt \rfloor}{nt} = t \lim_{x\to\infty}\frac{\lfloor x \rfloor}{x} = t \cdot 1 = t,$$

$$\lim_{n\to\infty}\left[\left(1 - \frac{\lambda}{n}\right)^n\right]^{\frac{\lfloor nt \rfloor}{n}} = \left[\lim_{n\to\infty}\left(1 - \frac{\lambda}{n}\right)^n\right]^{\lim_{n\to\infty}\frac{\lfloor nt \rfloor}{n}} = \left(e^{-\lambda}\right)^t = e^{-\lambda t}.$$

□

Exemple 4.3.3.

Veiem un exemple on, de forma natural, apareix la distribució exponencial com a límit de distribucions geomètriques.

Una nit d'estiu contemplem el firmament i anotem el temps (en minuts, amb tots el decimals que calgui) que triguem en observar una estrella fugaç, que anomenarem T i serà una variable aleatòria contínua. Si no sabem quin tipus de distribució contínua pot ser adient per modelar T, podríem començar per treballar amb models de variables aleatòries discretes.

Dividirem cada minut en n parts iguals (si $n = 60$ tindrem els 60 segons de cada minut) i definirem, per a $i \geq 1$,

$$Y_i^n = \begin{cases} 1 & \text{si a la part } i\text{-èsima apareix alguna estrella fugaç,} \\ 0 & \text{en cas contrari.} \end{cases}$$

Suposarem que Y_i^n, $i \geq 1$, són variables aleatòries independents Bernoulli de paràmetre

$$p_n = \text{“probabilitat de veure una estrella fugaç} \\ \text{en un interval de temps de durada } (1/n) \text{ minuts”.}$$

Anomenem X_n al nombre d'intervals de durada $(1/n)$ minuts que hem d'esperar fins a veure la primera estrella fugaç (dit d'una altra manera, X_n és l'índex i tal que Y_i^n és igual a 1 per primera vegada). Aleshores,

$$X_n \sim \text{Geomètrica}(p_n),\ \mathbb{E}(X_n) = \frac{1}{p_n}.$$

Observeu que X_n/n és el nombre de minuts (amb precisió de $(1/n)$ minuts) que triguem en observar la primera estrella fugaç. Per tant X_n/n és una aproximació de T, i aquesta aproximació serà tant millor com més gran sigui n. Com que

$$\mathbb{E}\left(\frac{X_n}{n}\right) = \frac{1}{np_n},$$

és raonable suposar que els paràmetres p_n s'han de triar de forma que

$$\mu = \mathbb{E}(T) = \frac{1}{np_n} \Leftrightarrow p_n = \frac{1}{n\mu} = \frac{\lambda}{n},$$

on $\lambda = 1/\mu$. D'aquesta manera, la funció de distribució de T hauria de ser el límit de les funcions de distribució de les X_n/n, quan n tendeix cap a infinit. Aleshores, l'Exemple 4.3.2 ens diu que la funció de distribució de T ha de ser

$$F_T(t) = \lim_{n\to\infty} \left[1 - \left(1 - \frac{\lambda}{n}\right)^{\lfloor nt \rfloor}\right] = 1 - e^{-\lambda t}.$$

Diem que T segueix una distribució exponencial de paràmetre λ. Aquesta és una manera natural d'introduir la distribució exponencial.

$\square$

Proposició 4.3.2. *Siguin $X, X_1, X_2, \ldots$ una successió de variables aleatorias discretes amb valors enters no negatius. Aleshores $X_n \xrightarrow{D} X$ quan $n \to \infty$. si i només si*

$$\Pr(X_n = k) \longrightarrow \Pr(X = k),$$

per a tot k enter no negatiu quan $n \to \infty$.

La demostració es bassa en les següents igualtats, certes per a variables aleatòries discretes amb valors enters no negatius:

$$\Pr(X = 0) = F_X(0), \Pr(X = k) = F_X(k) - F_X(k-1), \text{ per a tot } k \geq 1,$$

$$\Pr(X_n = 0) = F_{X_n}(0), \Pr(X_n = k) = F_{X_n}(k) - F_{X_n}(k-1), \text{ per a tot } k \geq 1,$$

$$F_X(k) = \sum_{j=0}^{k} \Pr(X = j), \text{ per a tot } k \geq 1,$$

$$F_{X_n}(k) = \sum_{j=0}^{k} \Pr(X_n = j), \text{ per a tot } k \geq 1.$$

Exemple 4.3.4.

Siguin $Y_n \sim B(n, p_n = \frac{\lambda}{n})$, per a tot n natural, $n \geq \lambda$. Sigui $Y \sim \text{Poisson}(\lambda)$. Aleshores $Y_n \xrightarrow{D} Y$.

Ho veurem veient que hi ha convergència de les funcions de probabilitat de Y_n cap a la de Y. Per a qualsevol enter no negatiu k tenim que

$$\lim_{n \longrightarrow \infty} \Pr(Y_n = k) = \lim_{n \longrightarrow \infty} \binom{n}{k} p^k (1-p)^{n-k} =$$

$$\lim_{n \longrightarrow \infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(\frac{n-\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} \lim_{n \longrightarrow \infty} \frac{n!}{(n-k)!} \frac{1}{n^k} \frac{(n-\lambda)^{n-k}}{n^{n-k}}$$

$$= \frac{\lambda^k}{k!} \lim_{n \longrightarrow \infty} \frac{n!}{(n-k)!} \frac{1}{n^n} (n-\lambda)^{n-k} = \frac{\lambda^k}{k!} \lim_{n \longrightarrow \infty} \frac{n!}{(n-k)!} \frac{(n-\lambda)^n}{n^n} \frac{1}{(n-\lambda)^k}$$

$$= \frac{\lambda^k}{k!} \lim_{n \longrightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \lim_{n \longrightarrow \infty} \frac{n(n-1)(n-2)\ldots 1}{(n-k)(n-k-1)\ldots 1} \frac{1}{(n-\lambda)^k}$$

(el primer límit és $e^{-\lambda}$)

$$= \frac{\lambda^k}{k!} e^{-\lambda} \lim_{n \longrightarrow \infty} n(n-1)(n-2)\ldots(n-k+1) \frac{1}{(n-\lambda)^k}$$

(observem que els primers factors dins del límit es comporten com n^k)

$$= \frac{\lambda^k}{k!}e^{-\lambda} \lim_{n\longrightarrow\infty} n^k \frac{1}{(n-\lambda)^k} = \frac{\lambda^k}{k!}e^{-\lambda} \left(\lim_{n\longrightarrow\infty} \frac{n}{n-\lambda} \right)^k = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Veiem doncs que calculant el límit arribem a l'expressió de la funció de probabilitat d'una variable aleatòria amb llei Poisson amb paràmetre λ. Per tant

$$\lim_{n\longrightarrow\infty} \Pr(Y_n = k) = \Pr(Y = k)$$

com volíem provar.

L'aplicació pràctica d'aquest resultat és la següent. Sigui $X \sim B(n,p)$, $\lambda = np$ i $Y \sim$ Poisson(λ). Si np és gran (diguem-ne $np > 1$) i p és petit (diguem-ne $p < 0.1$), aleshores

$$\Pr(X = k) \approx \Pr(Y = k), \text{ per tot } k \in \mathbb{N} \cup \{0\}.$$

$\square$

Exemple 4.3.3, pàgina 178. Continuació. A l'exemple de les estrelles fugaces, sigui Y el nombre d'estrelles fugaces vistes durant un interval de durada igual a un minut. Per aproximar la distribució de probabilitat de Y, podem considerar

$$Y_n = \sum_{i=1}^{n} Y_i^n$$

que serà el nombre d'intervals de durada $(1/n)$ segons on hi ha estat vista alguna estrella fugaç. Observeu que $Y_n \sim B(n, p_n = (\lambda/n))$ i que la distribució de Y hauria de ser el límit de les distribucions de Y_n, quan n tendeix a infinit. Aleshores l'Exemple 4.3.4 ens diu que

$$Y \sim \text{Poisson}(\lambda).$$

La versió d'aquest exemple en temps continu s'anomena **Procès de Poisson**, i té les següents característiques:

- Els temps $T_1, \ldots, T_n, \ldots$ entre esdeveniments (aparicions d'estrelles fugaces, al'exemple) consecutius tenen distribució exponencial de paràmetre λ i són independents els uns dels altres.

- El nombre d'esdeveniments que es produeixen en un interval de longitud t segueix una Poisson de paràmetre λ. Si dos intervals tenen intersecció buida, les variables de Poisson que compten els esdeveniments succeïts a cadascun d'ells, són independents.

La versió discreta (temps entre esdeveniments geomètrics i nombre d'esdeveniments en un interval donat binomials) és una aproximació del Procès de Poisson, millor com més gran és n.

4.3.2 Altres tipus de convergències de variables aleatorias

Considerem una successió de variables aleatorias $\{X_n\}_{n\geq 0}$ i una altra variable aleatoria X que podria ser el límit (en algun sentit) de X_n quan n tendeix a infinit. Tant X_n, $n \geq 0$, com X són funcions definides en un mateix espai mostral Ω. Per tant la successió $\{X_n\}_{n\geq 0}$ és una successió de funcions, i el possible límit X és una altra funció. Aleshores les definicions dels diferents tipus de convergències de successions de funcions que s'estudien a l'anàlisi matemàtica també són aplicables a les successions de variables aleatorias. En veurem uns quants exemples d'aquests tipus de convergència que són rellevants en probabilitat i estadística.

Finalment val la pena emfatitzar quelcom que hem dit abans: tant X_n, $n \geq 0$, com X són funcions definides en un mateix espai mostral Ω. Per tant, quan l'experiment aleatori es realitza i un element $\omega \in \Omega$ n'apareix com a resultat, els valors de totes les variables aleatòries X_n, $n \geq 0$, i X, queden determinats i valen $X_n(\omega)$, $n \geq 0$, $X(\omega)$, respectivament. En aquest moment el que tenim és ja una successió de nombres reals $\{X_n(\omega)\}_{n\geq 0}$ i un altre nombre real $X(\omega)$ que podria ser el seu límit.

Definició 4.3.3. *Siguin $X, X_1, X_2, \ldots$ variables aleatorias definides en algun espai de probabilitat $(\Omega, \mathcal{A}, \Pr)$. Direm que:*

(a) $X_n \to X$ *quan* $n \to \infty$ QUASI SEGURAMENT *si*

$$\Pr\left(\{\omega \in \Omega : X_n(\omega) \to X(\omega) \text{ quan } n \to \infty\}\right) = 1,$$

i escriurem

$$X_n \xrightarrow{q.s.} X.$$

(b) $X_n \to X$ *quan* $n \to \infty$ EN MITJANA r-ÈSIMA, *amb* $r \geq 1$, *si* $\mathbb{E}(|X_n|^r) < \infty$ *per tot* n *i*

$$\mathbb{E}(|X_n - X|^r) \to 0 \text{ quan } n \to \infty,$$

i escriurem

$$X_n \xrightarrow{r} X.$$

(c) $X_n \to X$ *quan* $n \to \infty$ EN PROBABILITAT *si*

$$\Pr(|X_n - X| > \epsilon) \to 0 \text{ quan } n \to \infty \text{ per a tot } \epsilon > 0,$$

i escriurem

$$X_n \xrightarrow{P} X.$$

Observeu que les definicions d'aquests tipus de convergències de successions de variables aleatorias necessiten que tant X_n, $n \geq 0$, com X estiguin definides en un mateix espai de probabilitat $(\Omega, \mathcal{A}, \Pr)$ (la successió de nombres reals $X_n(\omega)$, $n \geq 0$, ha de quedar determinada en el mateix moment que el possible límit $X(\omega)$, $|X_n - X|^r$ ha de poder-se calcular).

Això no era necessari quan parlàvem de convergència en distribució de X_n, $n \geq 0$, cap a X, perquè en aquest cas els únics elements que es tenen en compte en la definició són les funcions de distribució F_{X_n}, $n \geq 0$, i F_X. Per treballar amb les funcions de distribució no cal saber en quin espai mostral estan definides les variables aleatorias. De fet podem treballar directament amb successions de funcions de distribució F_n, $n \geq 0$, que convergeixen a una funció de distribució límit F, sense fer esment a cap variable aleatòria que tingui alguna d'aquestes funcions de distribució.

Relacions entre els tipus de convergències

Teorema 4.3.4. *Es tenen les següents implicacions entre els tipus de convergències de successions de variables aleatorias:*

$$\begin{array}{ccl} (X_n \xrightarrow{q.s.} X) & \searrow & \\ & & (X_n \xrightarrow{P} X) \Rightarrow (X_n \xrightarrow{D} X) \\ (X_n \xrightarrow{r} X) & \nearrow & \end{array}$$

per qualsevol $r \geq 1$. *A més a més, si* $r > s \geq 1$ *es té que*

$$(X_n \xrightarrow{r} X) \Rightarrow (X_n \xrightarrow{s} X).$$

En general no es donen més implicacions.

Com a exemple del tipus d'arguments que es fan servir per provar el resultat anterior, veurem que la convergència en mitjana quadràtica implica la convergència en probabilitat. L'argument és similar al que es fa servir per provar la desigualtat de Txebixev. Per tot $\varepsilon > 0$,

$$\begin{aligned}
\mathbb{E}[(X_n - X)^2] &= E\left[(X_n - X)^2 \left(I_{|X_n - X| > \varepsilon} + I_{|X_n - X| \le \varepsilon}\right)\right] \\
&= E\left[(X_n - X)^2 I_{|X_n - X| > \varepsilon}\right] + E\left[(X_n - X)^2 I_{|X_n - X| \le \varepsilon}\right] \\
&\ge E\left[(X_n - X)^2 I_{|X_n - X| > \varepsilon}\right] \ge \varepsilon^2 E\left[I_{|X_n - X| > \varepsilon}\right] \ge \varepsilon^2 \Pr(|X_n - X| > \varepsilon).
\end{aligned}$$

Aleshores

$$\Pr(|X_n - X| > \varepsilon) \le \frac{\mathbb{E}[(X_n - X)^2]}{\varepsilon^2}$$

i, per tant, si $\mathbb{E}[(X_n - X)^2]$ tendeix cap a 0 quan n tendeix cap a infinit, també $\Pr(|X_n - X| > \varepsilon)$ tendirà cap a 0, per tot $\varepsilon > 0$. Tenim doncs que si X_n convergeix a X en mitjana quadràtica aleshores també convergeix en probabilitat.

Teorema 4.3.5. *(a) Si $X_n \xrightarrow{D} c$, on c és una constant, aleshores $X_n \xrightarrow{P} c$.*

(b) Si $X_n \xrightarrow{P} X$ i $\Pr(|X_n| \le k) = 1$ per a tot n i algún k, aleshores $X_n \xrightarrow{r} X$ per a tot $r \ge 1$.

(c) Si $P_n(\epsilon) = \Pr(|X_n - X| > \epsilon)$ satisfà que $\sum_n P_n(\epsilon) < \infty$ per a tot $\epsilon > 0$, aleshores $X_n \xrightarrow{q.s.} X$.

Teorema 4.3.6. *(a) Si $X_n \xrightarrow{q.s.} X$ i $g : \mathbb{R} \longrightarrow \mathbb{R}$ és una funció contínua, aleshores $g(X_n) \xrightarrow{q.s.} g(X)$.*

(b) Si $X_n \xrightarrow{P} X$ i $g : \mathbb{R} \longrightarrow \mathbb{R}$ és una funció contínua, aleshores $g(X_n) \xrightarrow{P} g(X)$.

(c) Si $X_n \xrightarrow{D} X$ i $g : \mathbb{R} \longrightarrow \mathbb{R}$ és una funció contínua, aleshores $g(X_n) \xrightarrow{D} g(X)$.

Teorema 4.3.7. *Sigui $X_1, \ldots, X_n, \ldots$ una successió de variables aleatorias amb Var$(X_n) > 0$, per tot $n \ge 1$. Sigui $a \in \mathbb{R}$ una constant. Si es verifica que*

(i) $\mathbb{E}(X_n) \longrightarrow a$ quan $n \to \infty$, i

(ii) Var$(X_n) \longrightarrow 0$ quan $n \to \infty$,

aleshores X_n canvergeix a a en mitjana quadràtica i, per tant, també en probabilitat.

La demostració d'aquest resultat és conseqüència directa d'aquesta descomposició:

$$\mathbb{E}[(X_n - a)^2] = \mathbb{E}[(X_n - \mathbb{E}(X_n) + \mathbb{E}(X_n) - a)^2]$$

$$= \mathbb{E}[(\{X_n - \mathbb{E}(X_n)\} + \{\mathbb{E}(X_n) - a\})^2]$$

$$= \mathbb{E}[(X_n - \mathbb{E}(X_n))^2 + (\mathbb{E}(X_n) - a)^2 + 2(X_n - \mathbb{E}(X_n))(\mathbb{E}(X_n) - a)]$$

$$= \mathbb{E}[(X_n - \mathbb{E}(X_n))^2] + (\mathbb{E}(X_n) - a)^2 + 2(\mathbb{E}(X_n) - a)\mathbb{E}[(X_n - \mathbb{E}(X_n))]$$

$$= \mathrm{Var}(X_n) + (\mathbb{E}(X_n) - a)^2.$$

Aquest últim resultat és fa servir habitualment en estadística per demostrar que un ESTIMADOR T_n (una funció de les primeres n dades observades) convergeix en probabilitat a un PARÀMETRE a (una constant que representa alguna característica de interès de la distribució de probabilitat que genera les dades observades). En aquest cas es diu que T_n és un estimador CONSISTENT (en probabilitat) del paràmetre a. A la quantitat $\{\mathbb{E}(T_n) - a\}$ se l'anomena BIAIX de T_n com estimador de a. Aleshores el resultat anterior es pot enunciar de la manera següent:
Si el biaix de T_n com estimador del paràmetre a i la seva variància tendeixen a 0 quan n tendeix cap a infinit, aleshores T_n és un estimador consistent del paràmetre a.

4.4 Lleis dels grans nombres

Teorema 4.4.1 (Llei dèbil dels grans nombres)**.** *Siguin $X_1, X_2, \ldots$ variables aleatorias independents idènticament distribuïdes amb $\mathbb{E}(X_1) = \mu$ i $V(X_1) = \sigma^2 < \infty$. Sigui $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$. Aleshores*

$$\bar{X}_n \xrightarrow{\mathrm{Pr}} \mu \text{ quan } n \longrightarrow \infty.$$

Per demostrar-ho farem servir la desigualtat de Txebixev:

$$\mathrm{Pr}\left(|\bar{X}_n - \mu| > \varepsilon\right) \leq \frac{\mathrm{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2/n}{\varepsilon^2} \longrightarrow 0$$

quan $n \longrightarrow \infty$, i es té el resultat desitjat.

Teorema 4.4.2 (Llei forta dels grans nombres)**.** *Siguin* $X_1, X_2, \ldots$ *variables aleatorias independents idènticament distribuïdes. Aleshores*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow \mu \text{ quasi segurament quan } n \longrightarrow \infty$$

per alguna constant μ*, si i només si* $\mathbb{E}(|X_1|) < \infty$*. En aquest cas* $\mu = \mathbb{E}(X_1)$*.*

Exemple 4.4.1.

En el cas de tenir n variables aleatorias independents de Bernoulli de paràmetre p, la seva mitjana mostral coincideix amb la proporció mostral d'èxits en n experiments,

$$\hat{p}_n = \bar{X}_n = \frac{S_n}{n}$$

i el valor esperat μ és la probabilitat d'èxit p. En aquest cas la Llei dels Grans Nombres diu que

$$\hat{p}_n \longrightarrow_n p$$

en distribució i quasi segurament, és a dir, la proporció mostral d'èxits en n experiments tendeix a la probabilitat d'èxit quan el nombre d'experiments tendeix cap a infinit.

Aquest resultat reforça la interpretació freqüentista de la probabilitat. □

4.5 Teorema del Límit Central

Les lleis dels grans nombres ens diuen que la mitjana mostral $\bar{X}_n$ tendeix cap al valor esperat $\mu = E(X)$ o, en altres paraules, que $(\bar{X}_n - \mu)$ tendeix cap a 0 quan n tendeix cap a infinit.

De les propietats de la mitjana mostral es dedueix que

$$E(\bar{X}_n - \mu) = 0, \;\; \mathrm{Var}(\bar{X}_n - \mu) = \frac{\sigma^2}{n} \to_n 0,$$

on $\sigma^2 = \mathrm{Var}(X)$.

Observeu que si multipliquem les variables aleatòries $(\bar{X}_n - \mu)$ (que tendeixen a 0) per les constants $\sqrt{n}$ (que tendeixen a infinit) obtenim

$$E\left(\sqrt{n}(\bar{X}_n - \mu)\right) = 0, \;\; \mathrm{Var}\left(\sqrt{n}(\bar{X}_n - \mu)\right) = n\frac{\sigma^2}{n} = \sigma^2$$

per a tot n. És a dir, sembla que les variables aleatorias $\sqrt{n}(\bar{X}_n - \mu)$ *s'estabilitzen* en el sentit que tenen esperança i variància constants en n. El següent teorema (un dels més importants en estadística) ens diu que aquesta estabilització va encara més enllà.

Teorema 4.5.1 (Teorema del Límit Central)**.** *Siguin $X_1, X_2, \ldots$ variables aleatorias independents idènticament distribuïdes amb $0 < V(X_i) = \sigma^2 < \infty$ i $E(X_i) = \mu$. Sigui $S_n = \sum_{i=1}^n X_i$, $n \geq 1$. Aleshores*

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{D} N(0,1) \text{ quan } n \longrightarrow \infty.$$

Alternativament, si definim $\bar{X}_n = S_n/n$ podem escriure que

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{D} N(0,1) \text{ quan } n \longrightarrow \infty.$$

En termes pràctics, quan n sigui *gran* podrem fer les aproximacions

$$S_n \approx N(n\mu, n\sigma^2), \quad \bar{X}_n \approx N\left(\mu, \frac{\sigma^2}{n}\right),$$

en el sentit d'aproximar les funcions de distribucions (sovint desconegudes) de S_n i $\bar{X}_n$, per les funcions de distribucions de les corresponent distribucions normals que les aproximen.

Sorgeix aleshores la pregunta de quan gran ha de ser n per poder fer les aproximacions que es deriven del Teorema del Límit Central. La resposta és que depèn de com és la distribució de X:

- Si la distribució de X és discreta n ha de ser més gran que si és contínua.
- Si la distribució de X és asimètrica n ha de ser més gran que si és simètrica.
- Per a distribucions de X semblants a la normal no cal que el valor de n sigui molt gran.

En general si $n \geq 30$ les aproximacions seran bones.

Exemple 4.5.1.

Siguin $X_i, \ldots, X_n$ variables aleatorias independents Bernoulli de paràmetre p. Sigui $\hat{p}_n = (1/n)\sum_{i=1}^n X_i$. El Teorema del Límit Central ens diu que

$$\frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{p(1-p)}} \xrightarrow{D} N(0,1) \text{ quan } n \longrightarrow \infty,$$

o que, a la pràctica, podem aproximar la distribució de $\hat{p}_n$ per una $N(\mu = p, \sigma^2 = p(1-p)/n)$.

$\square$

4.5.1 Aproximacions per la normal de distribucions discretes

Siguin $X_i, \ldots, X_n$ variables aleatorias independents Bernoulli de paràmetre p. Sigui $X = \sum_{i=1}^{n} X_i \sim B(n,p)$. Per tant, una $B(n,p)$ és una suma de variables aleatorias independents i el Teorema del Límit Central ens diu que la seva distribució serà aproximadament normal:

$$B(n,p) \approx N(\mu = np, \sigma^2 = np(1-p)).$$

El significat d'aquesta aproximació ens el dona el mateix Teorema del Límit Central:

$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{D} Z \sim N(0,1).$$

És a dir, les variables aleatòries binomials centrades i tipificades convergeixen en distribució a una normal estàndard Z:

$$\Pr\left(\frac{X_n - np}{\sqrt{np(1-p)}} \leq z\right) \underset{n\to\infty}{\longrightarrow} \Pr(Z \leq z), \text{ per a tot } z \in \mathbb{R}.$$

A la pràctica, si n és *gran* farem la següent aproximació: per a tot $x \in \mathbb{R}$

$$\Pr(X_n \leq x) = \Pr\left(\frac{X_n - np}{\sqrt{np(1-p)}} \leq \frac{x - np}{\sqrt{np(1-p)}}\right) \approx \Pr\left(Z \leq \frac{x - np}{\sqrt{np(1-p)}}\right).$$

També es fa servir la següent aproximació, que inclou *correcció per continuïtat*: per a tot $k \in \mathbb{Z}^+$

$$\begin{aligned}
\Pr(X_n = k) &= \Pr(X_n \leq k + 0.5) - \Pr(X_n \leq k - 0.5) \\
&\approx \Pr\left(\frac{k - 0.5 - np}{\sqrt{np(1-p)}} < Z \leq \frac{k + 0.5 - np}{\sqrt{np(1-p)}}\right) \\
&= \Phi\left(\frac{k + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k - 0.5 - np}{\sqrt{np(1-p)}}\right),
\end{aligned}$$

on Φ és la funció de distribució de la normal estàndard.

Les aproximacions anteriors són bones quan n sigui gran i p no estigui massa a prop de 0 o de 1. Com a regla general podem dir que l'aproximació de la binomial per la normal és bona si

$$\min\{np, n(1-p)\} > 5.$$

Observeu que un argument similar ens diu que la distribució binomial negativa de paràmetres r i p també es pot aproximar per la distribució normal.

També és possible aproximar la distribució de Poisson per la normal quan el paràmetre λ de la Poisson sigui *gran*. Aquest resultat es pot justificar pel Teorema del Límit Central i per una propietat de la distribució de Poisson que no hem demostrat en aquest curs:

Si $X_1, \ldots, X_n$ són variables aleatorias independents de Poisson amb paràmetres $\lambda_1, \ldots, \lambda_n$, respectivament, aleshores la seva suma també és Poisson:

$$X = \sum_{i=1}^{n} X_i \sim \mathit{Poisson}\left(\lambda = \sum_{i=1}^{n} \lambda_i\right).$$

Per tant, si $X \sim \text{Poisson}(\lambda)$ amb λ *gran*, podem pensar que

$$\lambda = \sum_{i=1}^{n} \frac{\lambda}{n}$$

per un valor *moderat* de n i un valor no massa petit λ/n. Per tant

$$X = \sum_{i=1}^{n} X_i, \text{ amb } X_i, \ldots, X_n \text{ v.a.i.i.d. Poisson}(\lambda/n),$$

i el Teorema del Límit Central ens diu que la seva distribució serà aproximadament normal:

$$\text{Poisson} \approx N(\mu = \lambda, \sigma^2 = \lambda).$$

Aquesta aproximació és tant millor quan més gran és λ. En la pràctica es demana $\lambda \geq 10$.

4.6 Convergències de sumes de variables aleatorias i d'altres transformacions

Teorema 4.6.1. *Siguin $X, X_1, X_2, \ldots, Y, Y_1, Y_2, \ldots,$ variables aleatorias definides en el mateix espai de probabilitat $(\Omega, \mathcal{A}, \Pr)$. Aleshores:*

(a) Si $X_n \xrightarrow{q.s.} X$ i $Y_n \xrightarrow{q.s.} Y$ aleshores $X_n + Y_n \xrightarrow{q.s.} X + Y$.

(b) Si $X_n \xrightarrow{r} X$ *i* $Y_n \xrightarrow{r} Y$ *aleshores* $X_n + Y_n \xrightarrow{r} X + Y$.

(c) Si $X_n \xrightarrow{P} X$ *i* $Y_n \xrightarrow{P} Y$ *aleshores* $X_n + Y_n \xrightarrow{P} X + Y$.

(d) En general no és cert que $X_n + Y_n \xrightarrow{D} X + Y$ *quan* $X_n \xrightarrow{D} X$ *i* $Y_n \xrightarrow{D} Y$.

Tot i que no farem una demostració formal, val la pena assenyalar el següent. Als tres primers apartats, per a qualsevol possible resultat $\omega \in \Omega$ la suma $X_n(\omega) + Y_n(\omega)$ és cada vegada més propera a $X(\omega) + Y(\omega)$ perquè, per separat, $X_n(\omega)$ i $Y_n(\omega)$ són cada vegada més propers a $X(\omega)$ i $Y(\omega)$, respectivament. Per contra, a l'apartat *(d)*, quan hi ha convergència en distribució, no tenim garantia de que $X_n(\omega)$ i $Y_n(\omega)$ siguin cada vegada més propers a $X(\omega)$ i $Y(\omega)$, respectivament. Aleshores no podem tenir cap control de què passa amb la suma $X_n(\omega) + Y_n(\omega)$. El següent exemple ens mostra un cas extrem d'això.

Exemple 4.6.1.
Siguin $V_n \sim B(n,p)$, $n \geq 1$, i $Z \sim N(0,1)$ definides en el mateix espai de probabilitat $(\Omega, \mathcal{A}, \Pr)$. Definim

$$X_n = Y_n = \frac{V_n - np}{\sqrt{np(1-p)}},\ n \geq 1,\ X = Z,\ Y = -Z.$$

Observeu que $Y = -Z$ també té distribució $N(0,1)$. Pel Teorema del Límit Central sabem que

$$X_n \xrightarrow{D} X,\ Y_n \xrightarrow{D} Y.$$

Però $X_n + Y_n = 2X_n$ no convergeix en distribució a $X + Y = Z + (-Z) = 0$ (de fet $2X_n$ convergeix en dsitribució a $2Z \sim N(0,4)$). $\square$

Teorema 4.6.2 (Teorema de Slutsky)**.** *Siguin* $X, X_1, X_2, ..., Y_1, Y_2, ...,$ *variables aleatorias definides en el mateix espai de probabilitat* $(\Omega, \mathcal{A}, \Pr)$. *Si* $X_n \longrightarrow X$ *en llei i* $Y_n \longrightarrow c$ *en probabilitat, on c és una constant, aleshores:*

1. $X_n + Y_n \longrightarrow X + c$ *en llei.*
2. $X_n Y_n \longrightarrow cX$ *en llei.*
3. $X_n / Y_n \longrightarrow X/c$ *en llei, si* $c \neq 0$.

4. Suposem que $X_n \longrightarrow 0$ en llei i $Y_n \longrightarrow Y$ en probabilitat. Sigui $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ una funció contínua. Aleshores

$$g(X_n, Y_n) \xrightarrow{P} g(0, Y) \text{ quan } n \longrightarrow \infty.$$

Corol·lari 4.6.3. *Si $Y_n \longrightarrow Y$ en probabilitat i g és una funció contínua, aleshores $g(Y_n) \longrightarrow g(Y)$ en probabilitat.*

Exemple 4.6.2.

Siguin $X_1, \ldots, X_n, \ldots$ v.a.i.i.d. segons X. Siguin $\mu = \mathbb{E}(X)$, $\sigma^2 = \mathrm{Var}(X)$,

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,\ S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

Quin és el comportament límit de

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n}?$$

Pel Teorema del Límit Central sabem que

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{D} Z \sim N(0, 1).$$

Per altra banda, veurem que $S_n^2 \xrightarrow{\mathrm{Pr}} \sigma^2$:

$$\begin{aligned} S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \\ &= \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 + \frac{n}{n-1}\bar{X}_n^2 - \frac{2}{n-1}\bar{X}_n\sum_{i=1}^{n} X_i \\ &= \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \frac{n}{n-1}\bar{X}_n^2 \\ &\xrightarrow{\mathrm{Pr}} 1 \cdot \mathbb{E}(X^2) - 1 \cdot \mu^2 = \sigma^2. \end{aligned}$$

Hem aplicat la Llei dels Grans Nombres, el Teorema de Slutski (apartats 1 i 2) i el seu Corol·lari.

Ara escrivim

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} = \frac{\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}}{\sqrt{\frac{S_n^2}{\sigma^2}}} \xrightarrow{D} \frac{Z}{1} = Z \sim N(0, 1).$$

Hem aplicat el Teorema de Slutski (apartat 3) i el seu Corol·lari. □

Exemple 4.6.3.
Veurem que les distribucions t de Student convergeixen en distribució a una normal estàndard quan el nombre de graus de llibertat tendeixen a infinit.

Recordem que T_r segueix una t de Student amb r graus de llibertat si i només si

$$T_r \sim \frac{X}{\sqrt{Y_r/r}}$$

on $X \sim N(0,1)$, $Y_r \sim \chi_r^2$, X i Y_r independents. Per altra banda

$$Y_r \sim \sum_{i=1}^{r} Z_i^2$$

on $Z_1, \ldots, Z_r$ són v.a.i.i.d. $N(0,1)$. Aleshores,

$$T_r \sim \frac{X}{\sqrt{(1/r)\sum_{i=1}^{r} Z_i^2}} \xrightarrow{D} \frac{X}{1} = X \sim N(0,1).$$

Hem aplicat la Llei dels Grans Nombres, el Teorema de Slutski (apartat 3) i el seu Corol·lari.

$\square$