TEMA XII: ESTIMACIÓN PUNTUAL

A estimación puntual consiste na obtención de valores para os parámetros da distribución dunha povoación a partir dunha mostra de esta.

Sexa x unha variábel aleatória en estudo cuia distribución $f(x,\theta)$ depende dun parámetro θ que desexamos coñecer. Estimamos o valor de θ a partir da mostra aleatória simple $(x_1,x_2,...,x_n)$ onde os x_i son variábeis independentes igualmente distribuídas.

Estimador de θ = función da mostra que estima o valor de θ

$$\hat{\theta}_n = H(\vec{x}) = H(x_1,x_2,...,x_n) \; [*]$$

(Estimación de θ = valor do estimador para unha mostra determinada)

Unha mostra de tamaño n é un vector aleatório n-dimensional (a priori descoñecemos os seus valores), daquela o estimador [*] é unha variábel aleatória con función de distribución aproximada para mostras grandes polo TCL ao ser as x_i v.a. independentes e igualmente distribuídas.

Distribucións das Características Mostrais

- Distribución mostral da média
- Distribución mostral da varianza mostral
- Distribución mostral dunha proporción
- Distribución mostral da diferéncia de médias
- Cuasivarianza mostral

•Distribución mostral da média

Sexa $(x_1,...,x_n)$ unha mostra da variábel aleatória x en estudo que segue unha distribución calquera de média μ e desviación típica σ.

$$\bar{x}=\sum_{i=1}^{n}\frac{x_i}{n}$$ (a média mostral $\bar{x}$ é un bo estimador da média povoacional μ)

$\bar{\theta}_n = \bar{x}$ é unha v.a. cuia distribución compre:

$\mathbf{E(\bar{x})} = 1/n\cdot\Sigma_{i=1}^{n}E(x_i) = 1/n\cdot\Sigma_{i=1}^{n}\mu = 1/n\cdot n\cdot\mu = \boldsymbol{\mu}$

a média das médias de várias mostras é a média povoacional

$\mathbf{Var(\bar{x})} = Var(1/n\cdot\Sigma_{i=1}^{n}x_i) = 1/n^2\cdot Var(\Sigma_{i=1}^{n}x_i) \underset{\text{v.a.indep.}}{=} 1/n^2\cdot\Sigma_{i=1}^{n}Var(x_i) \underset{x_i=\text{dist.q.povoación}}{=} 1/n^2\cdot\Sigma_{i=1}^{n}\sigma^2 = 1/n^2\cdot n\cdot\sigma^2 = \boldsymbol{\sigma^2/n}$

Para mostras pequenas ($n<30$): a distribución de $\bar{x}$ depende da distribución da povoación de xeito exacto: se a povoación é normal, a média mostral é normal. Para mostras grandes ($n\geq 30$), polo TCL a média mostral póde-se aproximar por unha distribución normal (resultado aproximado)

•Distribución mostral da varianza

Sexa $(x_1,...,x_n)$ unha mostra da variábel aleatória x en estudo que segue unha distribución calquera de média μ e desviación típica σ.

$$S_n^2=\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{n}$$ (a varianza mostral S_n^2 é un estimador da varianza povoacional σ^2)

$\bar{\theta}_n= S_n^2$ é unha v.a. cuia distribución compre:

$$\mathbf{E(S_n^2)}=\frac{1}{n}E\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)=\frac{\mathbf{(n-1)\cdot\sigma^2}}{\mathbf{n}}$$ non obtemos a varianza da povoación

o caso ideal seria ter a varianza da povoación

$$\mathbf{E\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)}=E\left(\sum_{i=1}^{n}(x_i-\mu+\mu-\bar{x})^2\right)=E\left[\sum_{i=1}^{n}(x_i-\mu)^2+\sum_{i=1}^{n}(\mu-\bar{x})^2+\sum_{i=1}^{n}(x_i-\mu)(\mu-\bar{x})\right]=$$

$$=E(\Sigma(x_i-\mu)^2)+E(\Sigma(\mu-\bar{x})^2)+E(\Sigma(x_i-\mu)(\mu-\bar{x}))=n\cdot\sigma^2+n\cdot\sigma^2/n+(-2)\cdot n\cdot\sigma^2/n=$$
$$=n\sigma^2-\sigma^2=\mathbf{(n-1)\sigma^2}$$

Acerca da distribución de S_n^2 de xeito exacto, depende da distribución de x; de xeito aproximado, achéga-se á normal pero como para isto é preciso ter mostras mui grandes, non se usa.

• Distribución mostral dunha proporción

Nunha povoación hai individuos de duas clases A e B en proporción p e q=1-p respeitivamente. Ten-se unha mostra da povoación. A proporción mostral é o estimador da proporción.

$\bar{p}_n=x_n/n$ (x_n é o número de A's na mostra) asignando $x_i=1$ se $x_i\in A$ e 0 se $x_i\in B$

$\bar{p}_n$ estimador de p

$\bar{p}_n=1/n\cdot\Sigma_{i=1}^{n}x_i$ é unha variábel aleatória cuia distribución compre:

E($\bar{p}_n$)=p (a média da proporción da mostra é a probabilidade da mostra)

porque $\Sigma_{i=1}^{n} x_i$ é unha suma de v.a. de Bernouilli, entón $\Sigma_{i=1}^{n} x_i=n\bar{p}_n$ é unha binomial B(n,p), entón $E(\Sigma_{i=1}^{n} x_i)=E(n\bar{p}_n)=np$ e como $E(n\bar{p}_n)=E(\bar{p}_n)\cdot n$ entón $E(\bar{p}_n)=np/n=p$

Var($\bar{p}_n$)=pq/n

porque como $n\bar{p}_n\in B(n,p)$, $Var(n\bar{p}_n)=npq$, e como $Var(n\bar{p}_n)=n^2\,Var(\bar{p}_n)$, $Var(\bar{p}_n)=npq/n^2=pq/n$

acerca da distribución de $\bar{p}_n$

$n\bar{p}_n\in B(n,p)\approx N(np,npq)\Rightarrow \bar{p}_n\approx N(p,pq/n)$

e tipificando $(\bar{p}_n-p)/\sqrt{(pq/n)}\in N(0,1)$ aproximadamente

• **Distribución mostral da diferéncia de médias**

Sexa x unha povoación de média μ_x e desviación típica σ_x
Sexa y unha povoación de média μ_y e desviación típica σ_y
Sexa unha mostra $(x_1,...,x_n)$ de x e $(y_1,...,y_m)$ de y
A diferéncia de médias mostrais é o estimador da diferéncia de médias
$\theta=\mu_x-\mu_y$

$\bar{\theta}=\bar{x}-\bar{y}=1/n\cdot\Sigma_{i=1}^{n}x_i+1/m\cdot\Sigma_{i=1}^{m}y_i$

E($\bar{\theta}$)$=E(\bar{x}-\bar{y})=E(\bar{x})-E(\bar{y})=\boldsymbol{\mu_x-\mu_y}$

Var($\bar{\theta}$)=$\boldsymbol{\sigma_x^2/n+\sigma_y^2/m}$ porque $Var(\bar{x}-\bar{y})\underset{\substack{\uparrow\\ \text{indep.}}}{=}Var(\bar{x})-Var(\bar{y})$

acerca da distribución de $\bar{\theta}$

$\bar{\theta}\approx N(\mu_x-\mu_y,\sigma_x^2/n+\sigma_y^2/m)$, tipificando $(\bar{\theta}-E(\bar{\theta}))/\sqrt{Var(\bar{\theta})}\approx N(0,1)$

ousexa $\dfrac{(\bar{x}-\bar{y})-(\mu_x-\mu_y)}{\sqrt{[\sigma_x^2/n+\sigma_y^2/m]}}\approx N(0,1)$

•**Cuasivarianza mostral**

$(x_1,x_2,...,x_n)$

$\sigma^2\rightarrow\quad S^2=\sum_{n=1}^{n}\dfrac{(x_i-\bar{x})^2}{n}\Rightarrow E(S^2)=\sigma^2$

$\mu\rightarrow\quad S_n^2=\dfrac{\Sigma(x_i-\mu)^2}{n}$ (se descoñecemos σ^2 pero coñecemos μ)

Distribución da Média e da Varianza mostral en povoacións Normais

Média mostral

Sexa $x\sim N(\mu,\sigma)$

$\bar{x}=\sum_{i=1}^{n}x_i/n\underset{\uparrow}{\Rightarrow}\bar{x}\sim N(\mu,\sigma/\sqrt{n})$

aditividade da distribución normal

$x_i\sim N(\mu,\sigma)\Rightarrow\dfrac{x_i-\mu}{\sigma}\sim N(0,1)$

$\Sigma\left(\dfrac{x_i-\mu}{\sigma}\right)^2\sim X_n^2$

$\|$

$nS_n^2/\sigma^2\sim X_n^2$

Distribución da Cuasivarianza mostral

a cuasivarianza mostral defíne-se $\overline{S}_n^2 = \dfrac{n}{n-1} S_n^2 = \dfrac{1}{n-1} \Sigma_{i=1}^{n}(x_i - \overline{x})^2$

Se x ~ N(μ,σ) → $(x_1,x_2,...,x_n)$ mostra aleatória simple de x ⇒ $\dfrac{(n-1)S_n^2}{\sigma^2} \sim X_{n-1}^2$

$E\left(\dfrac{(n-1)S^2}{\sigma^2}\right) = n-1 \qquad \Rightarrow E(S^2)=\sigma^2$

$Var\left(\dfrac{(n-1)S^2}{\sigma^2}\right) = 2(n-1) \qquad \Rightarrow Var(S^2)=2\sigma^4/(n-1)$

Se a povoación de partida non é normal, entón:

se o tamaño da mostra é suficientemente grande ⇒ $\overline{x}=\Sigma\ x_i/n \Rightarrow \overline{x} \sim N(\mu,\sigma/\sqrt{n})$
para a varianza: también se pode aproximar a unha Normal, pero ésta é mui lenta e non se soe usar

Un estimador dun parámetro pode ser mellor ou peor, xa que é unha variábel aleatória e non sempre proporcionará unha boa apreciación da característica que desexa estimar. Por isto defínen-se unhas propriedades dos estimadores que nos permiten estabelecer comparacións entre diversos estimadores dun parámetro e poder escoller segundo nos interese. Estas propriedades son:

Propriedades dos estimadores

- insesgadez
- eficacia
- consisténcia

[1] **INSESGADEZ** Sexa x unha variábel aleatória tal que a sua distribución é coñecida salvo por un parámetro que a caracteriza, θ. Sexa $(x_1,x_2,...,x_n)$ unha mostra aleatória simple e $\overline{\theta}_n=H(x_1,x_2,...,x_n)$ un estimador deste parámetro θ.

Diremos que $\overline{\theta}_n$ é un estimador <u>insesgado</u> ou <u>centrado</u> de θ
se e só se $E(\overline{\theta}_n)=\theta$

Definimos o **sesgo** de $\overline{\theta}_n$ como $\boldsymbol{\theta - E(\overline{\theta}_n)}$

média de várias mostras

A insesgadez non é imprescindíbel para a bonanza do estimador.

Un estimador é asintoticamente insesgado se $\lim(n\to\infty)\ E(\overline{\theta}_n)=\theta$

Ex.: A média mostral $\overline{x}$ é un estimador insesgado de μ a média da povoación porque $E(\overline{x})=\mu$
Ex.: A varianza mostral S_n^2 non é un estimador insesgado de σ^2 porque $E(S_n^2)=(n-1)/n\cdot\sigma^2$
pero S_n^2 é un estimador asintoticamente insesgado porque $\lim(n\to\infty)\ \text{sesgo}(S_n^2)=0$
Ex.: A cuasivarianza mostral S^2 é un estimador insesgado de σ^2 porque $E(S^2)=\sigma^2$

Pode haber vários estimadores insesgados dun mesmo parámetro.

$\bar\theta_n^1$ $\bar\theta_n^2$ θ $\bar\theta_n^1$ $\bar\theta_n^2$
+—+——+——+—+

$\bar\theta_n$ é un estimador centrado/insesgado $E(\bar\theta_n)=\theta$
pero é malo porque sofre dunha grande varianza

Vemos que necesitamos unha medida que conxugue as duas características para a bonanza dun estimador: que tanto a varianza como o sesgo sexan pequenos.

[2] **EFICÁCIA**

Error Cuadrático Meio **$ECM(\bar\theta_n)=Var(\bar\theta_n)+Sesgo(\bar\theta_n)^2$**

$ECM(\bar\theta_n)=E(\bar\theta_n-\theta)^2=E(\bar\theta_n-E\bar\theta_n+E\bar\theta_n-\theta)^2=E(\bar\theta_n-E\bar\theta_n)^2+E(E\bar\theta_n-\theta)^2$
$+E(2(\bar\theta_n-E\bar\theta_n)(E\bar\theta_n-\theta))=Var(\bar\theta_n)+Sesgo(\bar\theta_n)^2+E(2(\bar\theta_n-E\bar\theta_n)(E\bar\theta_n-\theta))=$
$=Var(\bar\theta_n)+Sesgo(\bar\theta_n)^2+(E\bar\theta_n-\theta)2(E\bar\theta_n-E\bar\theta_n)=Var(\bar\theta_n)+Sesgo(\bar\theta_n)^2$

entón un estimador será mellor (*mais eficaz*) cando menor sexa o ECM
Procuramos entre os estimadores insesgados o mais eficaz:
sendo $\bar\theta_n$ insesgado $ECM(\bar\theta_n)=Var(\bar\theta_n)$, def.: $Eficacia(\bar\theta_n)=1/Var(\bar\theta_n)$

[3] **CONSISTÉNCIA**

cando non se pode obter estimadores insesgados de alta eficacia pero se dispón de <u>mostras grandes</u> esíxe-lle aos estimadores que sexan consistentes:

$$\mathbf{\lim_{n\to\infty} ECM(\bar\theta_n)}=\lim_{n\to\infty} Var(\bar\theta_n)+Sesgo(\bar\theta_n)^2=\lim_{n\to\infty} Var(\bar\theta_n)+(E(\bar\theta_n)-\theta)^2=\mathbf{0}$$

para o cal $\lim_{n\to\infty} Var(\bar\theta_n)=0$ e $\lim_{n\to\infty} E(\bar\theta_n)=0$

▪ <u>Exercício</u> [XII.1] *Demostrar que calquera combinación lineal de estimadores centrados da forma $\Sigma_n^{i=1}\lambda_i T_i$ para un parámetro T é tamén centrado se $\Sigma_n^{i=1}\lambda_i=1$*

$$E(\Sigma_n^{i=1}\lambda_i T_i)=\underbrace{\Sigma_n^{i=1}[\lambda_i\cdot E(T_i)]}_{\uparrow}=\underbrace{\Sigma_n^{i=1}[\lambda_i\cdot T]}_{\uparrow}=\underbrace{[\Sigma_n^{i=1}\lambda_i]\cdot T=T}_{\uparrow}$$

a esperanza é un operador lineal — os T_i son operadores insesgados — hipótese dada

---

▪ <u>Exercício</u> [XII.2] *Obter un estimador centrado para p nunha distribución binomial e calcular o seu erro meio cuadrático. É consistente?*

Sexa x unha variábel aleatória en estudo, $x \in B(k,p)$

Observamos k individuos con probabilidade p de ter unha característica, e miramos cantos teñen a característica x_i

$\vec{x}=(x_1,x_2,...,x_n)$

Exemplo:

x="nº de pezas defectuosas nunha caixa"
k=12 pezas por caixa

n=20 caixas
p=probabilidade de que a peza sexa defectuosa
$\vec{x}=(2,1,0,1,...,2)$
2=nº defectuosas na caixa 1
1=nº defectuosas na caixa 2
...
$\hat{p}$=(nº pezas defectuosas)/(nº pezas observadas)=
=(2+1+0+1+...+2)/(n·k)=(2+1+0+1+...+2)/(20·12)

$x\in B(k,p) \Rightarrow E(x)=kp$
$Var(x)=kpq$

$\underline{\Sigma_{i=1}^{n} x_i} \in B(nk,p) \approx N(nkp,\sqrt{(nkpq/n^2k^2)})$

$\uparrow$ aditividade da binomial

$\uparrow$
$y=\Sigma x_i$
$\alpha=nk$
$E(y/\alpha)=Ey/\alpha$
$dt(y/\alpha)=\sqrt{Var(y/\alpha)}=\sqrt{[Var(y)/\alpha^2]}=1/\alpha\cdot dt(y)$

$\hat{p}_n=\mathbf{\Sigma x_i/nk} \approx N(p,\sqrt{(pq/nk)})$

<u>propriedades</u>

sesgo = $\hat{p}_n$ = p-E$\hat{p}_n$ = p-p = 0 é un estimador insesgado

Var $\hat{p}_n$ = pq/nk

$ECM(\hat{p}_n)=sesgo^2+var=pq/nk \le 1/(4nk)=1/(48n)$

$$\left.\begin{array}{l} p\in(0,1) \\ q=1-p \end{array}\right\} pq\le 0.25$$

$1/(48n) \xrightarrow{n\to\infty} 0$ é un estimador consistente

---

▪ <u>Exercício</u> [XII.3] *Para estimar a média dunha povoación considéra-se o estimador a·`x. Atopar o valor de a que minimiza o erro cuadrático meio da estimación.*

$ECM=Sesgo(a\,\bar{x})^2+Var(a\,\bar{x})$
$=[E(a\,\bar{x})-\mu]^2+a^2Var(\bar{x})=$
$=[a\cdot E(\bar{x})-\mu]^2+a^2\cdot\sigma^2/n=$
$=[a\cdot\mu-\mu]^2+a^2\cdot\sigma^2/n$
$\Rightarrow$
$\phi(a)=[a\cdot\mu-\mu]^2+a^2\cdot\sigma^2/n=a^2\mu^2+\mu^2-2a\mu^2+a^2\sigma^2/n$
$\phi'(a)=2a\mu^2-2\mu^2+2a\sigma^2/n=0$
$\Leftrightarrow a(2\mu^2+2\sigma^2/n)=2\mu^2$
$\Leftrightarrow a=2\mu^2/(2\mu^2+2\sigma^2/n)=\mu^2/(\mu^2+\sigma^2/n)$

▪ Exercício [XII.4] *Unha firma de mercadotécnia envia un cuestionário a 1,000 residentes de certo subúrbio dunha cidade para determinar as suas preferéncias como compradores. Dos 1,000 residentes 80 responden o cuestionário. O anterior constitúe unha mostra aleatória? Dicer as vantaxes deste procedemento para obter unha mostra aleatória.*

Tráta-se dunha mostra aleatória, ainda que non obtida por mostreo aleatório simple. A maior vantaxe é que é un sistema barato.

▪ Exercício [XII.5] *Nunha pranta de armado automotriz seleccionarán-se 50 dos primeiros 1,000 automóbiles dun novo modelo para ser inspeccionados polo departamento de control de calidade. O xerente da pranta decide inspeccionar un automóbil cada vez que rematen de armarse 20. Este proceso dará como resultado unha mostra aleatória?*

Tráta-se de mostro sistemático: tóma-se un primeiro elemento e despois os seguintes a intervalos iguais.

▪ Exercício [XII.6] *Para un determinado nivel de ingresos o Departamento de Facenda sabe que as cantidades declaradas por conceito de deduccións médicas (x_1), contribucións caritativas (x_2), e gastos vários (x_3) son variábeis independentes, normais con médias \$400, \$800, e \$100 e desviacións típicas \$100, \$250, \$40 respectivamente.*

i) Cal é a probabilidade de que a cantidade total declarada por conceito destas tres deduccións, non sexa maior de \$1,600?

ii) Se unha persoa con este nivel de ingresos declara por conceito destas deduccións un total \$2,100, que tan probábel é ter unha cantidade igual ou maior a este monto baixo as condicións dadas?

i)

Como as x_i son v.a.i. polo TCL:

$S_n=x_1+x_2+x_3 \approx N(E(S_n),\sqrt{Var(S_n)})$

$E(S_n)=\Sigma\, \overline{x}_i=400+800+100=1,300$

$\sqrt{Var(S_n)}=\Sigma\, s_i^2=100^2+250^2+40^2=10,000+62,500+1,600=74,100$

$S_n \in N(1,300,272.21315)$

$P(S_n\leq 1,600)=P(z\leq 1.1020775)=F(1.10)=0.8643$

ii)

$P(S_n\geq 2,100)=P(z\geq 2.9388735)=1-P(z<2.94)=1-0.998359=0.001641$

▪ Exercício [XII.7] *Nun experimento binomial obsérvan-se x éxitos en n ensaios independentes. Propóñen-se as duas seguintes funcións como estimadores do parámetro p: $T_i=x/n$; $T_2=(x+1)/(n+2)$.*

a) Obter e comparar os erros meios cuadráticos para T_1 e T_2.

b) Demostrar que T_1 é un estimador consistente de p.

a)

$x\in B(n,p)\Rightarrow E(x)=np;\ Var(x)=npq$

$ECM(T)=Sesgo(T)^2+Var(T)=[E(T)-p]^2+Var(T)$

$ECM(T_1)=[E(T_1)-p]^2+Var(T_1)=$

$=[E(x/n)-p]^2+Var(x/n)=$

$=[1/n\cdot E(x)-p]^2+Var(x/n)=$

$=[1/n\cdot np-p]^2+1/n^2\cdot Var(x)=$

$=1/n^2\cdot npq=$

$=p(1-p)/n$

$ECM(T_2)=[E(T_2)-p]^2+Var(T_2)=$

$=[E((x+1)/(n+2))-p]^2+Var(x/(n+2)+1/(n+2))=$
$=[1/(n+2)\cdot E(x)+1/(n+2)-p]^2+1/(n+2)^2\cdot Var(x)=$
$=[(np+1)/(n+2)-p]^2+np(1-p)/(n+2)^2=$
$=[(np+1-pn-2p)/(n+2)]^2+np(1-p)/(n+2)^2=$
$=(1-2p)^2/(n+2)^2+np(1-p)/(n+2)^2=$
$=[(1-2p)^2+np(1-p)]/(n+2)^2$

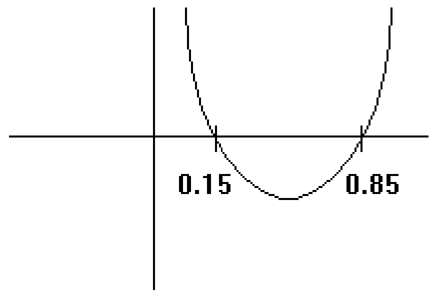
$$¿\ \frac{p(1-p)}{n} < \frac{[(1-2p)^2+np(1-p)]}{(n+2)^2}\ ?$$

$\Leftrightarrow n(0.25+2p^2-2p) < p(1-p)$

Analisamos as posíbeis solucións:

- caso trivial é $0.24+2p^2-2p=0 \Leftrightarrow p=0.15$ ou 0.85

  no intervalo $(0.15,0.85)$ T_2 é mellor que T_1



- en caso contrário depende de n: se $n\to\infty$, T_1 é mellor que T_2

b)

$\lim_{n\to\infty} ECM(T_1)=\lim_{n\to\infty} p(1-p)/n=0 \Rightarrow T_1$ é un estimador consistente de p

---

- <u>Exercício</u> [XII.8] *Sexa x_1, x_2, x_3, x_4 unha mostra aleatória de tamaño 4 dunha povoación de distribución expoñencial con parámetro q descoñecido.*
  - *a) Das seguintes funcións, cales son estimadores insesgados de q?*
    - $T_1=1/6(x_1+x_2)+1/3(x_3+x_4)$
    - $T_2=(x_1+2x_2+3x_3+4x_4)/5$
    - $T_3=(x_1+x_2+x_3+x_4)/4$
  - *b) De entre os estimadores insesgados de q, determinar cal é o que ten a varianza mais pequena.*
  - *c) Cales son as eficácias relativas dos demais estimadores insesgados con respeito ao que ten a varianza mais pequena?*

a)

$x \in Exp(\theta) \Rightarrow E(\theta)=\theta;\ Var(\theta)=\theta$

$E(T_1)=E(1/6(x_1+x_2)+1/3(x_3+x_4))=1/6E(x_1)+1/6E(x_2)+1/3E(x_3)+1/3E(x_4)=$
$=(2/6+2/3)\theta=\theta \Rightarrow T_1$ é un estimador insesgado de θ

$E(T_2)=E((x_1+2x_2+3x_3+4x_4)/5)=0.2\cdot(10)\cdot\theta=2\cdot\theta$
$\Rightarrow T_2$ é un estimador sesgado de θ

$E(T_3)=E((x_1+x_2+x_3+x_4)/4)=0.25(E(x_1)+E(x_2)+E(x_3)+E(x_4))=\theta$
$\Rightarrow T_3$ é un estimador insesgado de θ

b)

$Var(T_1)=Var(1/6x_1+1/6x_2+1/3x_3+1/3x_4)=$
$=1/36Var(x_1)+1/36Var(x_2)+1/9Var(x_3)+1/9Var(x_4)=$
$=(1/18+2/9)\cdot\theta=5/18\cdot\theta=0.28\cdot\theta$

$Var(T_3)=Var((x_1+x_2+x_3+x_4)/4)=$
$=1/16\cdot 4\cdot\theta=1/4\cdot\theta=0.25\cdot\theta$

$\Rightarrow$

$Var(T_1)>Var(T_3)$

c)

A eficiéncia (ou eficácia) dun estimador é maior canto menor sexa o ECM (ousexa, eficácia=1/ECM). Como ECM=sesgo^2+varianza e nos casos de T_1 e T_3 son insesgados (sesgo=0), daquela medimos a eficácia pola varianza: do aparatado anterior, T_3 é un estimador mais eficiente porque ten menos varianza.

Píden-nos a eficácia de T_1 respeito de T_3, iso é o cociente:

$$\frac{\mathrm{efic}(T_1)}{\mathrm{efic}(T_3)} = \frac{1/\mathrm{ECM}(T_1)}{1/\mathrm{ECM}(T_3)} = \frac{1/\mathrm{Var}(T_1)}{1/\mathrm{Var}(T_3)} = \frac{\mathrm{Var}(T_3)}{\mathrm{Var}(T_1)} = 0.25/0.28 = 0.89 \approx 90\%$$

▪ Exercicio [XII.9] *Sexa x_1, x_2, x_3, x_4, x_5 unha mostra aleatória dunha povoación cuia distribución é normal. Considerar como estimadores da média povoacional* m *as seguintes funcións:*

$T_1=(x_1+x_2+x_3+x_4+x_5)/5$

$T_2=(x_1+x_2+2x_3+x_4+x_5)/6$

Identificar o estimador de varianza mais pequena.