Guía 9: apuntamos a buscar estimadores de un parámetro de una distribución

Es decir, $X \sim F_\theta$, ¿θ?

m.a. = muestra aleatoria es $\underline{X} = (X_1, \dots X_n)$ tq $X_i \overset{iid}{\sim} F_\theta$

Buscamos un estimador de θ que sea una VA función de $\underline{X}$

Algunas def previas:

$X \sim f_\theta$ $\longrightarrow$ es la densidad de X si es VAC o es la $p(x)$ si X es VAD

<u>Función de Verosimilitud</u>: Dado $\underline{X}$, $L(\theta) = f(x_1, x_2, \dots x_n)$

$$= \prod_{i=1}^{n} f_\theta(x_i)$$

<u>Familias exponenciales</u>: $X \sim F_\theta$

Decimos que la distribución es una familia exponencial

si $f_\theta(x) = A(\theta)\, e^{\sum_{i=1}^{k} c_i(\theta) r_i(x)} \cdot h(x)$ $\forall x$, $\theta = (\theta_1, \dots, \theta_k)$,

$k = \#$ parámetros

9.4 Mostrar que las siguientes familias de distribuciones son familias exponenciales a 1 parámetro: (**a**) Bernoulli(p); (**b**) Pascal$(4,p)$; (**c**) Poisson(λ); (**d**) Exponencial(λ).

(d)

$$X \sim \mathcal{E}(\lambda) \Rightarrow f_\lambda(x) = \underbrace{\lambda}_{A(\lambda)} e^{-\lambda x} \overbrace{\mathbb{1}\{x > 0\}}^{h(x)} \quad , K = 1$$

$c(\lambda) = -\lambda$, $r_1(x) = x$ ó $c(\lambda) = \lambda$, $r_1(x) = -x$ ó $c(\lambda) = 2\lambda$, $r_1(x) = \frac{-x}{2}$

(e)

$$X \sim Geo(p) \Rightarrow f_p(x) = (1-p)^{x-1} p \, \mathbb{1}\{x \in \mathbb{N}\}$$

$$= p\, e^{\underbrace{x \log(1-p)}} e^{-\log(1-p)} \underbrace{\mathbb{1}\{x \in \mathbb{N}\}}_{h(x)}$$

$A(p)$

$c(p) = \log(1-p)$, $r(x) = x$

(hay mas opciones, podria por ejemplo $r(x) = (x-1)$ ...

$$f_p(x) = \underbrace{p}_{A(p)} e^{(x-1)\log(1-p)} \underbrace{\mathbb{1}\{x \in \mathbb{N}\}}_{h(x)}$$

$c(p) = \log(1-p)$
$r(x) = (x-1)$

(c)

$X \sim \mathcal{P}oi(\mu)$

$c(\mu) = \log \mu$

$r(x) = x$

$$f_\mu(x) = \frac{\mu^x}{x!} e^{-\mu} \mathbb{1}\{x \in \mathbb{N}_0\} = \underbrace{\frac{1}{x!} \mathbb{1}\{x \in \mathbb{N}_0\}}_{h(x)} e^{x \log \mu} \underbrace{e^{-\mu}}_{A(\mu)}$$

(f)

$X \sim \mathcal{U}(0, \theta)$ no es familia regular!!

El soporte depende de $\theta \Rightarrow$ no es familia exponencial

(g)

$X \sim \mathcal{N}(\mu, \sigma^2)$

$$f_{(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi}\,(\sigma^2)^{1/2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \mathbb{1}\{x \in \mathbb{R}\}$$

$$= \frac{1}{\sqrt{2\pi}\,(\sigma^2)^{1/2}} e^{-\frac{1}{2\sigma^2}x^2 + \frac{1}{\sigma^2}\mu x - \frac{1}{2\sigma^2}\mu^2} \underbrace{\mathbb{1}\{x \in \mathbb{R}\}}_{h(x)=1}$$

$$A(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\,(\sigma^2)^{1/2}}\, e^{-\frac{\mu^2}{2\sigma^2}} \qquad C_1(\mu, \sigma^2) = \frac{-1}{2\sigma^2}$$

$$C_2(\mu, \sigma^2) = \frac{\mu}{\sigma^2}$$

$$\sigma_1(x) = x^2$$

$$\sigma_2(x) = x$$

Tarea: Ver si son flia exponencial todas las distribuciones que usamos hasta ahora en particular $N(\mu_0, \sigma^2)$ y $N(\mu, \sigma_0^2)$

conocido

Estadístico $\longrightarrow$ Es una función de la m.a: $T(\underline{X})$

Estadístico suficiente para θ $\longrightarrow$ Es $T(\underline{X})$ tq $\underline{X} \mid T(\underline{X}) = t$ no depende de θ

¿Como se prueba que T es suficiente?

$\longrightarrow$ Por def (ejercicio 9.1)

$\longrightarrow$ Teorema de factorización $f_\theta(\underline{x}) = f(x_1, \dots, x_n)$

$$= g(T(\underline{x}), \theta) \cdot h(\underline{x}) \ \forall \underline{x}$$

$\longrightarrow$ Para familias exponenciales:

si $X \sim F_\theta$, $\theta = (\theta_1, \dots, \theta_k)$

$$\Rightarrow T(\underset{\sim}{X}) = \left(\sum_{i=1}^{n} r_1(X_i), \sum_{i=1}^{n} r_2(X_i), \dots, \sum_{i=1}^{n} r_k(X_i) \right)$$

es un estadístico suficiente para θ

9.2 Sea $\mathbf{X}_n$ una muestra aleatoria de tamaño n de una distribución con densidad

$$f_\theta(x) = e^{-(x-\theta)} \mathbf{1}\{x > \theta\}.$$

Verificar que $T = \min(X_1, \dots, X_n)$ es un estadístico suficiente para θ.

$\mathbb{1}\{x > \theta\}$ — Depende de $\theta \longrightarrow$ no es familia exponencial

$\Longrightarrow$ Uso teorema de factorización

$$f_\theta(\underline{x}) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} e^{-|x_i-\theta|}\, \mathbb{1}\{x_i > \theta\}$$

$$= \underbrace{\prod_{i=1}^{n} e^{-x_i}}_{h(x)} \; \underbrace{e^{n\theta} \underbrace{\prod_{i=1}^{n} \mathbb{1}\{x_i > \theta\}}_{= \mathbb{1}\{\min x_i > \theta\}}}_{g(\min x_i,\, \theta)}$$

$\Rightarrow$ $\boxed{T(X) = \min\{X_i,\ i=1\ldots n\} \text{ es suf para } \theta}$

$\downarrow$

en el estadístico es X, no x, ya que es una función de las VA

	θ parametro	$c(\theta)$	$r(x)$	Estadístico suf
$\mathcal{E}(\lambda)$	λ	$\dots$		$\sum_{i=1}^{n} X_i$
$\vdots$				
$N(\mu,\sigma^2)$	(μ,σ^2)			$\left(\sum_{i=1}^{n} X_i^2, \sum_{i=1}^{n} X_i\right)$
$\vdots$				

Método de máxima verosimilitud para hallar estimador

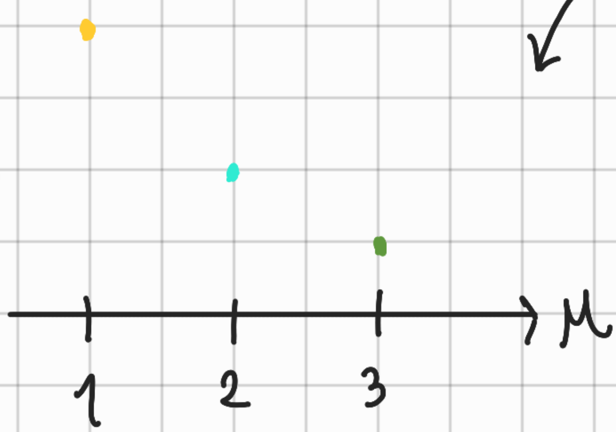
Idea: $X \sim N(\mu, 1)$ y μ puede ser 1, 2 o 3 $\Rightarrow$ Espacio paramétrico $\{1, 2, 3\}$

Si observo $x_1 = 1.2$ $(n=1)$ ¿Cuál será el μ estimado a partir de esta muestra?

$L(\mu)$ se llama función de verosimilitud $\Downarrow$ El μ estimado es $\mu = 1$

Dibujo $\underbrace{f_\mu}(1.2)$

Los valores de densidad de la m.a. observada



el μ más verosímil es el valor estimado de μ

$$L(\mu) = f_\mu(\underline{x}) = \prod_{i=1}^{n} f_\mu(x_i)$$

$\hat{\Theta}_{MV}$ es el estimador máximo verosímil de Θ si

$$\hat{\theta}_{MV} = \arg\max L(\theta, \underline{X}) = T(\underline{X}) \text{ es un estadístico}$$

9.6 La cantidad de accidentes de tránsito por semana en la intersección de Paseo Colón y Estados Unidos tiene una distribución de Poisson de media λ.

(**a**) Hallar el estimador de máxima verosimilitud de λ basado en una muestra aleatoria de la cantidad de accidentes durante n semanas. Mostrar que se trata de un estimador insesgado para λ y hallar la expresión de su error cuadrático medio.

(**b**) En una muestra de 100 semanas se observaron las frecuencias:

Cantidad de accidentes	0	1	2	3	4	5
Frecuencia	10	29	25	17	13	6

En virtud de la información muestral calcular el ~~estimador~~ Valor estimado de máxima verosimilitud de λ y estimar la probabilidad de que en la semana del 24 de diciembre de 2017 no ocurra ningún accidente en la mencionada esquina.

(a) X: # Accidentes en 1 semana $\sim P(\lambda)$

m.a. $\underline{X} = (X_1, \dots, X_n)$ busco EMV

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \mathbb{1}\{x_i \in \mathbb{N}_0\}$$

$$= \lambda^{\sum x_i} e^{-n\lambda} \prod_{i=1}^{n} \frac{\overbrace{\mathbb{1}\{x_i \in \mathbb{N}_0\}}^{\text{esto vale 1 ya que los } X_i \text{ son un número en los naturales}}}{x_i!}$$

Buscamos el máximo de $L(\lambda)$ en $\underbrace{\{\lambda > 0\}}$

Como no me dicen en qué espacio paramétrico uso el natural de la Poisson

Una estrategia cuando la familia es regular es maximizar $\log L(\lambda)$

En la indicadora no hay λ

$$\log L(\lambda) = \left(\sum_i x_i\right) \log \lambda - n\lambda + \log\left(\frac{\pi\{\dots\}}{x_i!}\right)$$

$$\frac{d}{d\lambda} \log L(\lambda) = \sum_1^n x_i \frac{1}{\lambda} - n = 0 \Longrightarrow \lambda = \frac{\sum_1^n x_i}{n} = \overline{x}$$

Pto crítico

Verifico que es máx:

$$\frac{d^2}{d\lambda^2} \log L = \underbrace{\sum x_i}_{\oplus} \underbrace{\left(-\frac{1}{\lambda^2}\right)}_{\ominus} < 0 \Rightarrow \text{Es máximo}$$

a la función de λ de x_i, la evalúo en el vector aleatorio

El estimador MV de λ es

$$\Longrightarrow \hat{\lambda}_{MV} = \frac{\sum_{n=1}^{\infty} X_i}{n} = \overline{X}$$

6

$0, 0, 0 \ldots \leftarrow 10$ veces

$1, 1, \ldots \leftarrow 29$ veces

Valor estimado de λ: $\hat{\lambda} = \bar{x} = \dfrac{\sum x_i}{100}$

$$= \frac{0 \cdot 10 + 1 \cdot 29 + 2 \cdot 25 + 3 \cdot 17 + 4 \cdot 13 + 5 \cdot 6}{100} = 2{,}12$$

Ahora me piden estimar $\mathbb{P}(X=0) = \dfrac{\lambda^0}{0!} e^{-\lambda} = \alpha(\lambda)$ función de λ

<u>Propiedad</u>: Principio de invarianza

si $\hat{\theta}$ es estimador MV de θ y $\alpha(\theta)$ es función biyectiva

$$\Rightarrow \widehat{\alpha(\theta)}_{MV} = \alpha(\hat{\theta}_{MV})$$

Obs: Aún cuando α no sea biy (iny) puede usarse ppio de invarianza

$$\Longrightarrow \widehat{P(X=0)} = e^{-\hat{\lambda}}$$ evaluando:

El valor estimado es: $\widehat{\mathbb{P}(X=0)} = e^{-2,12} = 0,12$

9.8 El tamaño, X (en GB), de ciertos archivos es una variable aleatoria cuya densidad es

$$f_\theta(x) = 3\theta^3 x^{-4} \mathbf{1}\{x \geqslant \theta\}, \qquad \theta > 0.$$ ← no es flia regular, no miro derivadas

(**a**) Hallar el estimador de máxima verosimilitud de θ basado en una muestra aleatoria de los tamaños de n archivos.

(**b**) Hallar expresiones para la esperanza y la varianza del estimador de máxima verosimilitud de θ.

(**c**) Mostrar que el estimador de máxima verosimilitud de θ converge en media cuadrática al verdadero valor de θ.

(a)

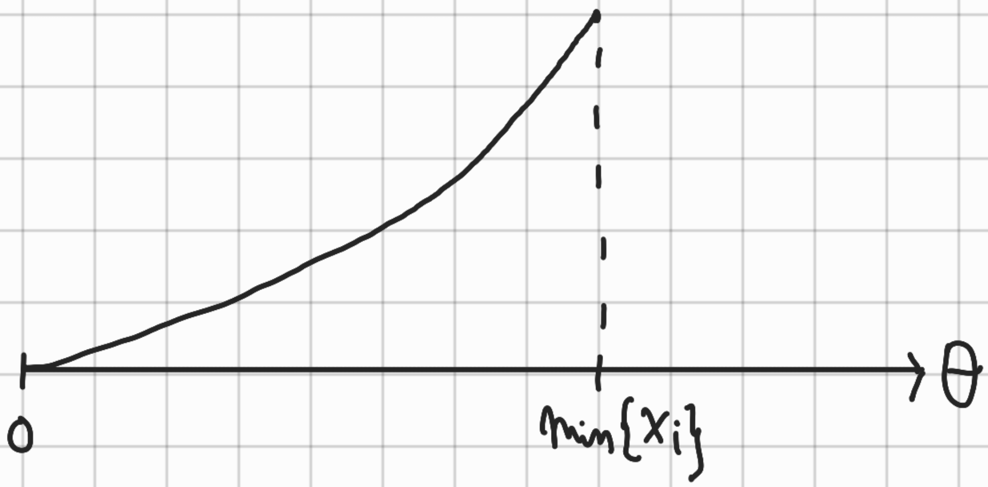
no es flia reg ← el sop depende de θ

no puedo usar log

$$L(\theta) = \prod_{i=3}^{n} 3\,\theta^3 x_i^{-4}\, \mathbb{1}\{x_i \geq \theta\}$$

$$= 3^n \theta^{3n} \underbrace{\prod_{i=1}^{n} x_i^{-4}\, \mathbb{1}\{x_i \geq \theta\}}_{\prod_{i=1}^{n} x_i^{-4} \cdot \underbrace{\prod_{i=1}^{n} \mathbb{1}\{x_i \geq \theta\}}_{\mathbb{1}\{\theta \leq \min x_i\}}}$$

$$L(\theta) = \underbrace{3^n}_{\text{cte}} \theta^{3n} \underbrace{\prod_{1}^{n} x_i^{-4}}_{\text{cte respecto de } \theta} \mathbb{1}\{\theta \leq \min x_i\}$$



$\Rightarrow$ El máximo de $L(\theta|x)$ se da en $\min\{x_i\}$

$$\Rightarrow \hat{\theta}_{MV} = \min\{X_i,\ i=1,\dots,n\}$$

Def: un estimador $\hat{\theta}$ es insesgado si $E(\hat{\theta}) = \theta$

Ej: dijimos que $X \sim Poi(\mu) \Rightarrow \hat{\mu} = \overline{X}$

Idea: el estimador es insesgado si la v.a. $\hat{\mu}$ da "alrededor de" lo que quiero estimar

En este caso $E(\hat{\mu}) = E[\overline{X}] = E\left(\frac{\sum_i X_i}{n}\right) = \dots = E(X_1) = \mu$

¡$\hat{\mu}$ es insesgado!

Se llama sesgo del estimador a $B(\hat{\theta}) = E[\hat{\theta}] - \theta$

Entonces, algo que podríamos pedirle a un estimador:

- "que de alrededor del parámetro" $\longrightarrow$ sesgo chico
- "que tenga baja variabilidad" $\longrightarrow$ $Var(\hat{\theta})$ chica
  (más concentrado)

Para controlar ambos se define:

Error cuadrático medio: $ECM(\hat{\theta}) = B(\hat{\theta})^2 + V(\hat{\theta})$

Volviendo al 9.6

Vimos que $\hat{\lambda} = \bar{X}$

$$\Rightarrow E[\hat{\lambda}] = \lambda \Rightarrow B(\hat{\lambda}) = \lambda - \lambda = 0 \longrightarrow \text{Es insesgado}$$

$$\Rightarrow V(\hat{\lambda}) = V(\bar{X}) = V\left(\frac{\sum X_i}{n}\right) = \frac{V(X_i)}{n} = \frac{\lambda}{n}$$

Entonces:

$$ECM(\hat{\lambda}) = 0 + \frac{\lambda}{n} = \frac{\lambda}{n}$$

Teorema:

Si $ECM(\hat{\theta}) \xrightarrow[n \to \infty]{} 0$ se dice que $\hat{\theta} \longrightarrow \theta$ en ECM

9.13

(**a**) Mostrar que la familia de distribuciones $\mathcal{N}(\mu, \sigma^2)$ puede expresarse en la forma

$$f_\theta(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \exp\left(\frac{\mu}{\sigma^2}x - \frac{1}{2\sigma^2}x^2\right),$$

donde $\theta = (\mu, \sigma^2)$.

(**b**) Sea $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n de la distribución $\mathcal{N}(\mu, \sigma^2)$. Hallar la expresión de la densidad conjunta y mostrar que $T = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ es un estadístico suficiente para θ.

(**c**) Sea $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$. Mostrar que $\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$ y deducir que $T' = (\bar{X}, S^2)$, donde

$$S^2 := \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2, \longleftarrow \text{Varianza muestral}$$

es un estadístico suficiente para θ.

(**d**) Hallar el estimador de máxima verosimilitud para θ basado en la muestra aleatoria $X_1, \ldots, X_n$.

(a) y (b) ya lo hicimos

(c)

Propiedad:

Si $T(\underline{X})$ es suficiente para θ y $T'(\underline{X}) = W(T(\underline{X}))$ con W inyectiva

$\Longrightarrow$ T' es suficiente para θ

(T' no es la devuelta)

$$T(\underline{X}) = \Big(\underbrace{\sum_i^n X_i}_{a} \,,\, \underbrace{\sum_i^n X_i^2}_{b} \Big)$$

$$T'(\underline{X}) = \left(\frac{\sum X_i}{n} \,,\, \frac{\sum X_i^2 - n\left(\frac{\sum X_i}{n}\right)^2}{n-1} \right)$$

Vemos que $T' = \left(\frac{a}{n} \,,\, \frac{b - a^2/n}{n-1} \right) = w(a,b) = (c,d)$

$$\begin{cases} c = a/n \\ d = \dfrac{b - a^2/n}{n-1} \end{cases} \rightsquigarrow \begin{cases} a = nc \\ b = (n-1)d + \dfrac{c^2}{n} \end{cases}$$

se ve que es inyectiva

$\Longrightarrow$ T' es suficiente para (μ, σ^2)

(d) $X \sim \mathcal{N}(\mu, \sigma^2)$; $\mu \in \mathbb{R}$, $\sigma^2 > 0$

$$\underbrace{L(\mu, \sigma^2)}_{} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,(\sigma^2)^{1/2}}\, e^{-\frac{1}{2\sigma^2}x_i^2 + \frac{\mu}{\sigma^2}x_i - \frac{\mu^2}{2\sigma^2}}$$

$$= \frac{1}{(\sqrt{2\pi})^n}\frac{1}{(\sigma^2)^{n/2}}\, e^{-\frac{n\mu^2}{2\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}\sum x_i^2 + \frac{\mu}{\sigma^2}\sum x_i}$$

cte

2 parametros $\Rightarrow$ maximizo función de 2 variables

Primero aplico el logaritmo:

$$\log L(\mu, \sigma^2) = \log \text{Cte} - \frac{n}{2}\log(\sigma^2) - \frac{n\mu^2}{2\sigma^2} - \frac{1}{2\sigma^2}\sum x_i^2 + \frac{\mu}{\sigma^2}\sum x_i$$

$$\begin{cases} \dfrac{\partial \log L}{\partial \mu} = -\dfrac{n}{2\sigma^2}2\mu + \dfrac{\sum x_i}{\sigma^2} = 0 \\[2ex] \dfrac{\partial \log L}{\partial \sigma^2} = -\dfrac{n}{2}\dfrac{1}{\sigma^2} + \dfrac{n\mu^2}{2(\sigma^2)^2} + \dfrac{\sum x_i^2}{2(\sigma^2)^2} - \dfrac{\mu \sum x_i}{(\sigma^2)^2} = 0 \end{cases}$$

De la primera ecuación $\rightsquigarrow \mu = \dfrac{\sum x_i}{n}$

Ahora voy a la segunda

$$\rightarrow \frac{-n\sigma^2 + n\mu^2 + \sum x_i^2 - 2\mu \sum x_i}{2(\sigma^2)^2} = 0$$

$$-n\sigma^2 + n\frac{\left(\sum x_i\right)^2}{n^2} + \sum x_i^2 - 2\frac{\sum x_i}{n}\sum x_i = 0$$

$$n\sigma^2 = \frac{\left(\sum x_i\right)^2}{n} + \sum x_i^2 - \frac{2}{n}\left(\sum x_i\right)^2 = \sum x_i^2 - \frac{n^2}{n}\underbrace{\left(\frac{\sum x_i}{n}\right)^2}_{\bar{X}}$$

$$\sigma^2 = \frac{1}{n}\left[\sum x_i^2 - n\bar{x}^2\right] = \frac{\sum (x_i - \bar{x})^2}{n}$$

$$\Rightarrow \hat{\theta}_{MV} = (\hat{\mu}, \hat{\sigma}^2) = \left(\bar{X}, \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}\right) = \left(\bar{X}, \frac{(n-1)S^2}{n}\right)$$

Falta ver las 2^{das} derivadas!! Hacerlo!! esto es lo que prueba (recién) que tenemos el EMV

Que puedo decir del sesgo en este caso $X \sim \mathcal{N}(\mu, \sigma^2)$

$$E(\hat{\mu}) = E(\bar{X}) = \mu$$

no es insesgado!

$$E(\hat{\sigma}^2) = E\left(\frac{(n-1)}{n} S^2\right) = \frac{n-1}{n} E(S^2) = \frac{n-1}{n} \sigma^2$$

En el video se prueba que $E(S^2) = \sigma^2$

9.11 En una mesa electoral votaron 129 ciudadanos. Se extrajeron (sin reposición) 7 sobres al azar de la urna, se examinaron y resultó que el candidato verde obtuvo exactamente 3 votos. Estimar por máxima verosimilitud la cantidad de votos por el candidato verde que había en la urna.

X: # votos verdes en 7 sobres $\sim$ Hipergeométrica $(129, d, 7)$

Dato (muestra observada, solo 1, no voy a haber una productoria en $L(d)$)

$: X_1 = 3$

¿Qué es la verosimilitud? es la probabilidad "de lo que observo" como función de "d"

$$L(d) = \frac{\binom{d}{3}\binom{129-d}{4}}{\binom{129}{7}} \mathbb{1}\left\{\begin{array}{l} d \in \mathbb{N} \\ d \geq 3 \\ 129-d \geq 4 \end{array}\right\}$$

Truco para hayar max en una función discreta $\frac{L(d+1)}{L(d)} \geq 1$

$\iff L$ crece