Departament d'Estadística i Investigació Operativa

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Probabilitat i variables aleatòries Annex amb exemples de servidors i connexions

Bloc A – Probabilitat i Estadística

2024

Índex

Exemple "1 Servidor i *n* connexions"

Un client es vol connectar amb un servidor remot mitjançant una xarxa de comunicacions. El procés consisteix en realitzar ***n intents de connexió*** *a la xarxa en un període determinat. Tenim èxit si, en algun intent, hem trobat un camí per la xarxa fins al servidor i el servidor està en marxa.*

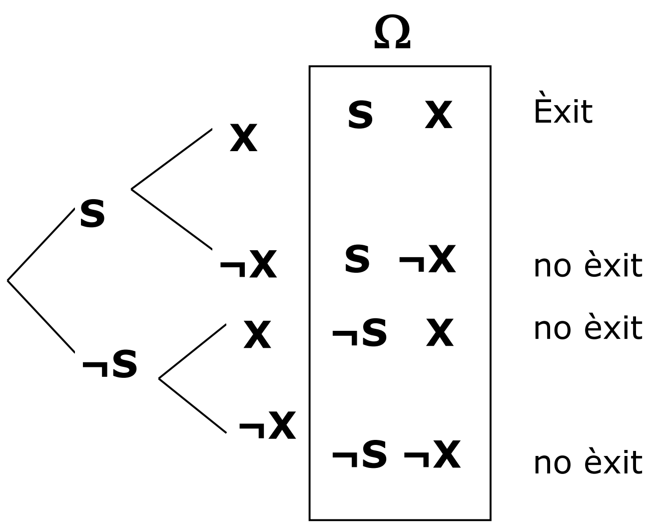
En primer lloc representarem l'experiència pels casos de 1 i 2 intents:

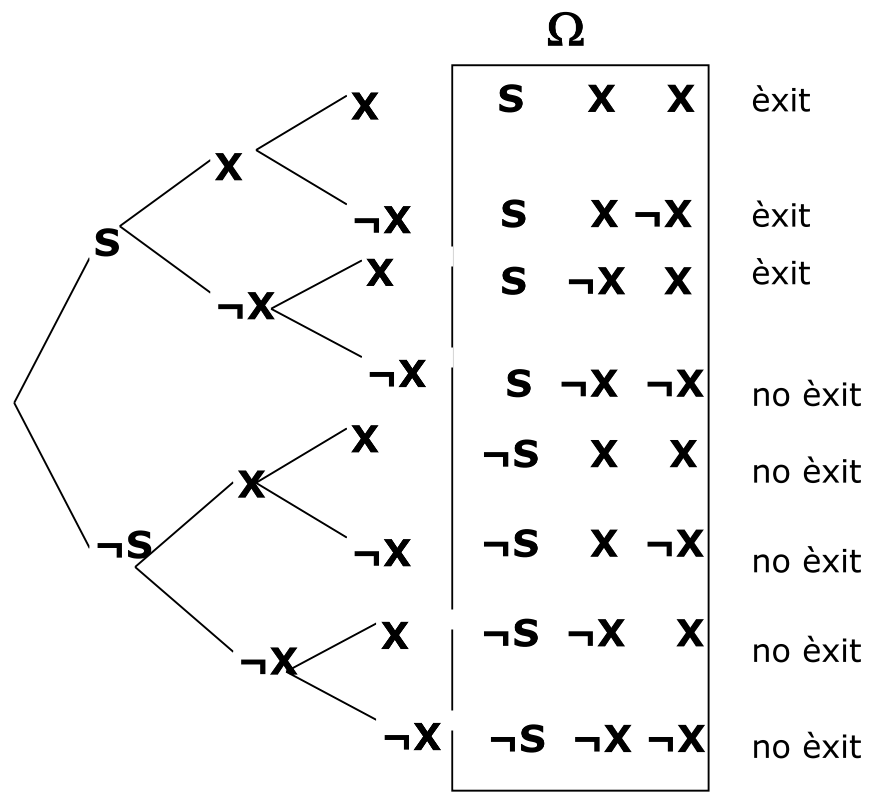
S = "el servidor està en marxa" (respon)
¬S = "no en marxa" (no respon)
X = "la petició del client ha trobat camí per la xarxa"
¬X = "no camí a la xarxa"

1) Representeu l'arbre pels casos 1 i 2 intents

Exemple "1 Servidor i *n* connexions"

Arbres de probabilitat





Source: (repàs de combinatòria: Lecture4 a Instructor Resources a http://www.janehorgan.com/)

Exemple “1 Servidor i *n* connexions”

Si en aquest exemple podem suposar:

- en n intents, el servidor no canvia d’estat
- els estats del servidor i de la xarxa són independents
- els n intents de connexió són independents uns d’altres
- el servidor falla, a l’atzar, 1 de cada 10 vegades: $P(\neg S)=p1=1/10$
- i la xarxa una de cada 5 vegades $P(\neg X)=p2=1/5$

Podem calcular la probabilitat d’èxit, és a dir, de contactar i poder treballar amb el servidor si només es realitza un intent, fent:

$$P(\text{el servidor funciona} \cap \text{la xarxa funciona}) = P(S \cap X) =$$

$$= P(S)\cdot P(X) = 9/10 \cdot 4/5 = 0.72$$

Independents

Exemple “1 Servidor i *n* connexions”

Si definim T_i com l’esdeveniment connectar en algun dels i intents, llavors podem calcular la probabilitat d’èxit si:

– només es realitza un intent

$$P(T_1) = P(S \cap X) = P(S) \cdot P(X) = (1 - p1) \cdot (1 - p2)$$

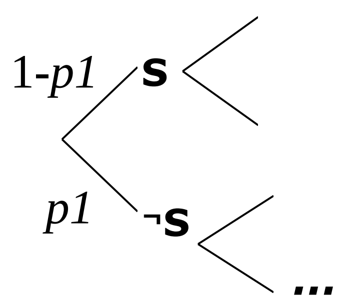
– es realitzen dos intents (1–Prob(“no èxit en 2 intents”))

$$P(T_2) = 1 - P(\neg T_2) = 1 - (p1 + (1 - p1) \cdot (p2)^2)$$

$$on \quad P(\neg T_2) = P(\overline{S}) + P(S \cap \overline{X} \cap \overline{X}) = p1 + (1 - p1) \cdot (p2)^2$$

– es realitzen n intents (1–Prob(“no èxit en n intents”))

$$P(T_n) = 1 - P(\neg T_n) = 1 - (p1 + (1 - p1) \cdot (p2)^n) \rightarrow_{n \to \infty} 1 - p1$$

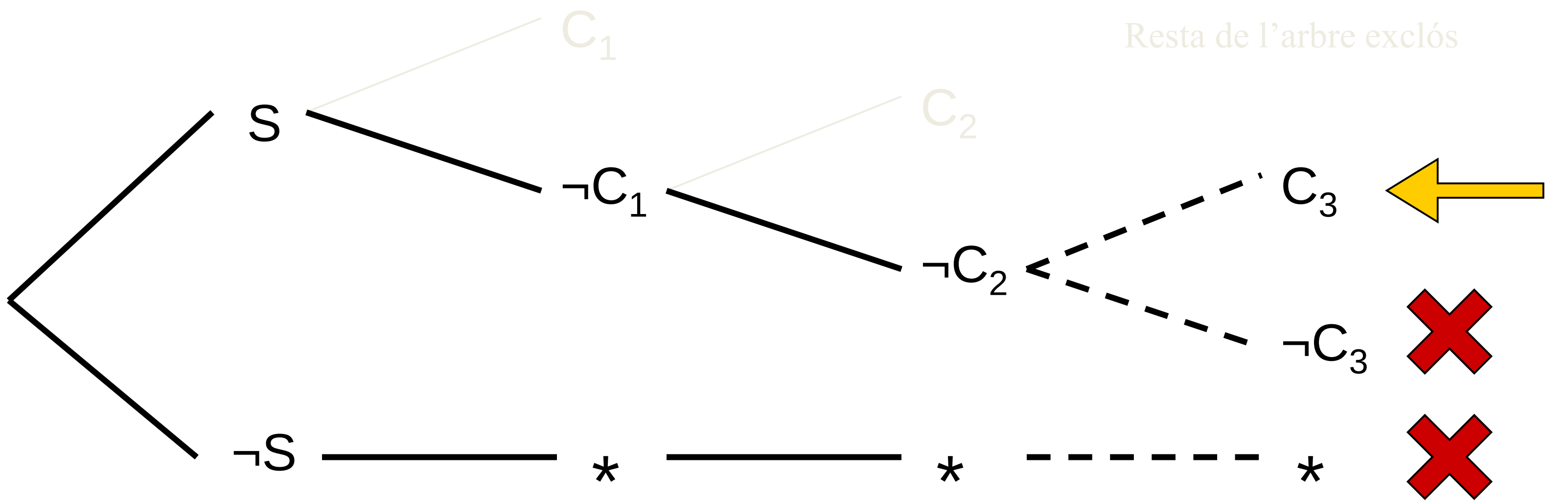


Si *n* augmenta, la probabilitat d’èxit tendeix a ***1-p1***

(tendeix a P(S), la probabilitat que el servidor funcioni, ja que si *n* és gran la probabilitat que en algun intent la xarxa no falli tendeix a 1)

Exemple "1 Servidor i *n* connexions"

Ara suposem que s'han realitzat dos intents de connexió sense èxit (no sabem si per causa de la xarxa o del servidor). ¿Quina és la probabilitat de connectar en un tercer intent?



Sabent que teníem definit T_i com l'esdeveniment "connectar en algun dels i intents", llavors $\neg T_2$ serà "no connectar en 2 intents". Així, cal calcular:

$$P(S \cap X_3 \mid \neg T_2) = \frac{P(S \cap X_3 \cap \neg T_2)}{P(\neg T_2)}$$

Exemple "1 Servidor i *n* connexions"

En primer lloc calculem la probabilitat del numerador. Com que:

$$S\cap X_3\cap\neg T_2 = S\cap X_3\cap [\neg S\cup(S\cap\neg X_1\cap\neg X_2)] = S\cap\neg X_1\cap\neg X_2 \cap X_3$$

llavors la seva probabilitat és:

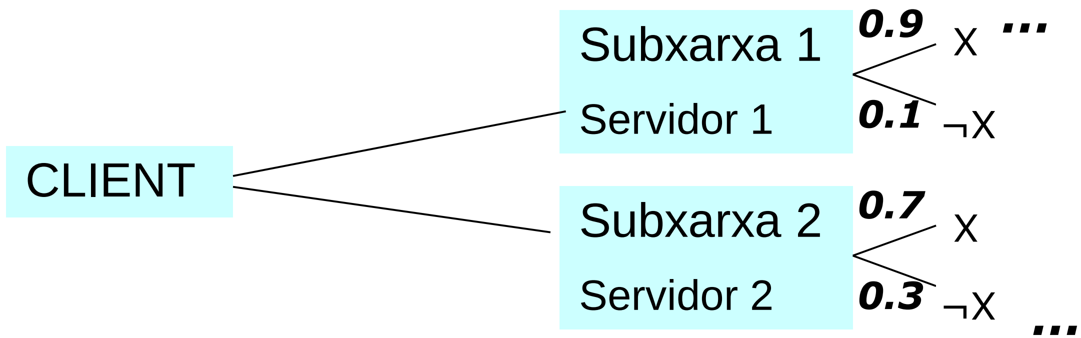
$$P(S\cap X_3\cap\neg T_2\ ) = 0.9\cdot 0.2\cdot 0.2\cdot 0.8 = 0.0288$$

Fallar en dos intents i tenir èxit en el tercer (amb el servidor en marxa) és un succés amb poques "possibilitats" (menys del 3%). I així, la probabilitat que busquem es compensa quan es compara amb el succés que condiciona ($\neg T_2$ que tampoc és molt freqüent):

$$P(S\cap X_3\mid\neg T_2) = \frac{P(S\cap X_3\cap\neg T_2)}{P(\neg T_2)} = \frac{P(S\cap X_3\cap\neg T_2)}{P\left[\neg S\cup(S\cap\neg X_1\cap\neg X_2)\right]} = \frac{0.0288}{0.136} = 0.2118$$

Aquesta probabilitat també la podem comparar amb la de connectar en un tercer intent sense tenir cap informació prèvia (probabilitat "bruta" de connectar en el tercer intent és $P(S \cap X_3)=0.72$). Per tant, no connectar en els dos intents previs és una mala senyal: ¡les possibilitats baixen del 72% a un 21%!

Exemple "2 Servidors i 2 connexions"

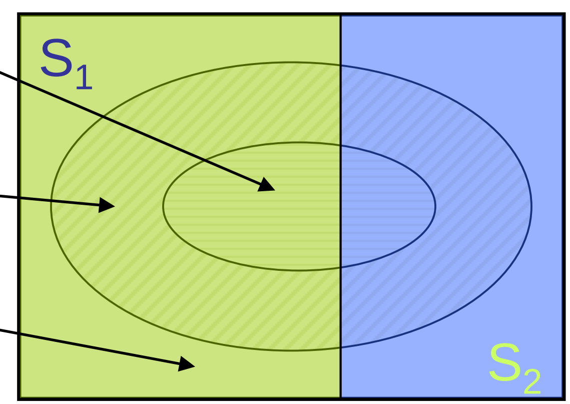


Ara considerem que tenim dues subxarxes amb dos servidors i en triem una o altra a l'atzar (50%). I llavors fem els intents sempre sobre la mateixa xarxa (en la primera, falla la connexió 1 de cada 10 cops; i en la segona, 3 de cada 10). Per a n=2 (2 intents de connexió). ¿Com calcular la probabilitat d'obtenir Y = 0, 1, 2 connexions? Assumeixi la independència que calgui.

conjunt de resultats associat a "**2 connexions**" (Y_2)

conjunt de resultats associat a "**1 connexió**" (Y_1)

conjunt de resultats associat a "**0 connexions**" (Y_0)



Exemple "2 Servidors i 2 connexions"

En aquest cas podem calcular fàcilment les probabilitats de Y_0, Y_1 o Y_2 condicionades pel servidor:

$P(Y_0 \mid S_1) = 0.1 \cdot 0.1 = 0.01$
$P(Y_0 \mid S_2) = 0.3 \cdot 0.3 = 0.09$
$P(Y_1 \mid S_1) = 0.1 \cdot 0.9 + 0.9 \cdot 0.1 = 0.18$
$P(Y_1 \mid S_2) = 0.3 \cdot 0.7 + 0.7 \cdot 0.3 = 0.42$
$P(Y_2 \mid S_1) = 0.9 \cdot 0.9 = 0.81$
$P(Y_2 \mid S_2) = 0.7 \cdot 0.7 = 0.49$

Quines probabilitats han de sumar 1?

I com que Y_0, Y_1 o Y_2 poden ser expressats com a unió de conjunts disjunts (ja que $\{S_1, S_2\}$ és una *partició*):

LPT

$$Y_i = (Y_i \cap S_1) \cup (Y_i \cap S_2); \quad \Rightarrow P(Y_i) = P(Y_i \cap S_1) + P(Y_i \cap S_2) = P(Y_i \mid S_1) \cdot P(S_1) + P(Y_i \mid S_2) \cdot P(S_2)$$

Calcula:

$P(Y_0) \rightarrow$ **0** connexions, 1 de cada 20 vegades

$P(Y_1) \rightarrow$ **1** connexió, 3 de cada 10 vegades

$P(Y_2) \rightarrow$ **2** connexions, 13 de cada 20 vegades

Exemple "2 Servidors i 2 connexions"

I ara suposant que s'han aconseguit k connexions, **amb quina probabilitat hem estat atesos pel servidor *i*?**

Agafant el nombre de connexions com una partició i els dos servidors com una altra partició podem aplicar:

$$P(A_i | B_k) = \frac{P(B_k | A_i) \cdot P(A_i)}{\sum_j P(B_k | A_j) \cdot P(A_j)}$$

Calculeu les probabilitats a posteriori

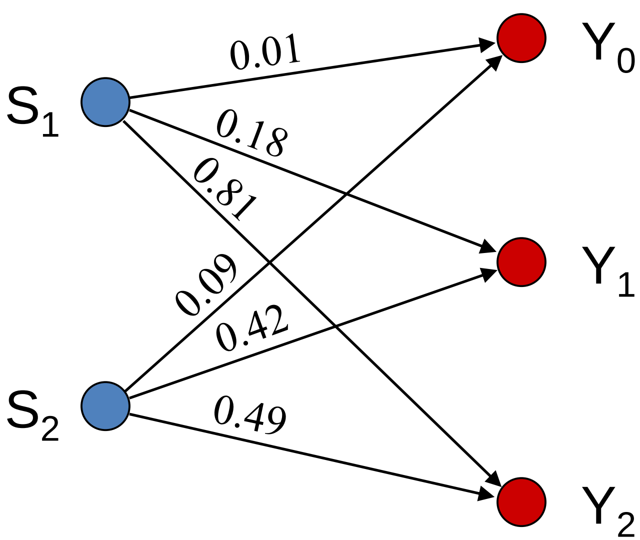
A priori	*i=1*	*i=2*	**P(Yk)**		A posteriori	*i=1*	*i=2*
P(Y_0 | Si)	0.01	0.09	0.05		**P(Si | Y_0)**		
P(Y_1 | Si)	0.18	0.42	0.30		**P(Si | Y_1)**		
P(Y_2 | Si)	0.81	0.49	0.65		**P(Si | Y_2)**		
P(Si)	0.5	0.5					

Si s'han aconseguit dues connexions, creiem que hem utilitzat el primer servidor amb probabilitat **.....**

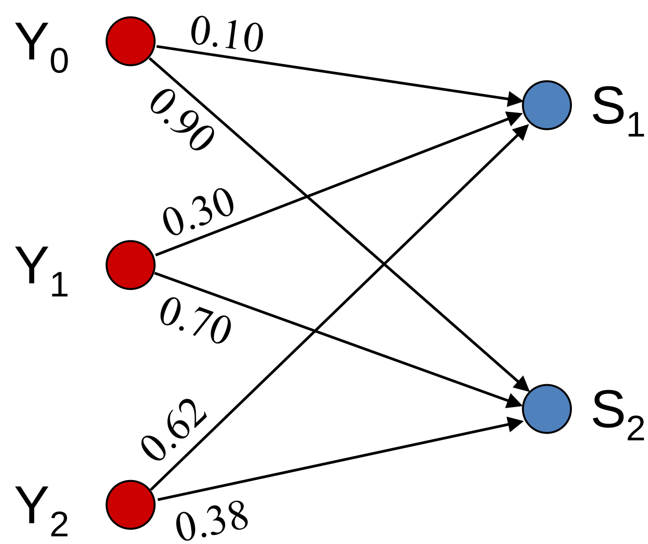
Exemple "2 Servidors i 2 connexions". Prob posteriori

El teorema de Bayes transforma unes probabilitats condicionades en unes altres

$P(Y_k \mid S_i)$



$P(S_i \mid Y_k)$



Coneixent el servidor utilitzat, calculem la probabilitat d'obtenir un número determinat de connexions.

Sabent el número de connexions aconseguides, calculem la probabilitat d'haver utilitzat determinat servidor.

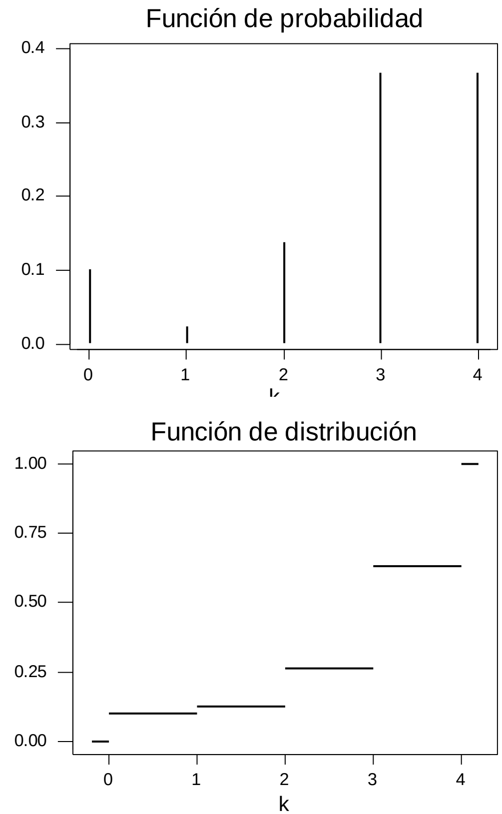
Exemple "1 Servidor i 4 connexions" amb VAD

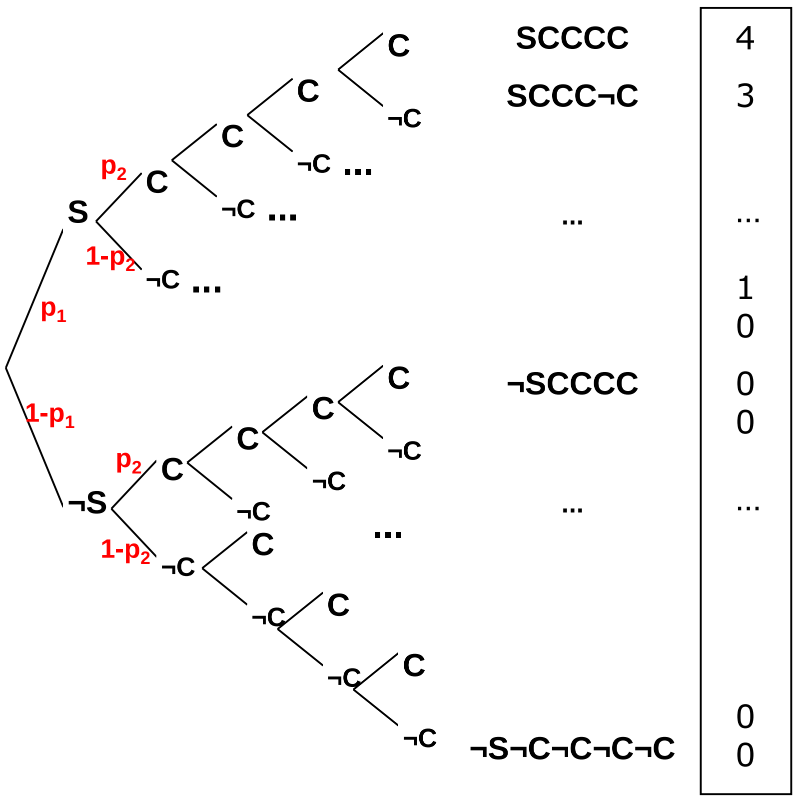
Reprenem l'exemple del servidor i la xarxa. Ara amb 4 intents de connexió, probabilitat de funcionar el servidor p_1=0.9 i probabilitat de funcionar la xarxa p_2= 0.8. Definim la VAD X = "nº de connexions aconseguides"

$P(X=0)= 1- p_1+ p_1\cdot(1- p_2)^4 \rightarrow \quad P(X=k)= p_1\cdot p_2^k\cdot(1- p_2)^{4-k}\binom{4}{k} \quad k>0$

k	$P_X(k)$
0	0.10144
1	0.02304
2	0.13824
3	0.36864
4	0.36864

k	$F_X(k)$
$(-\infty,0)$	0.00000
[0,1)	0.10144
[1,2)	0.12448
[2,3)	0.26272
[3,4)	0.63136
$[4,+\infty)$	1.00000

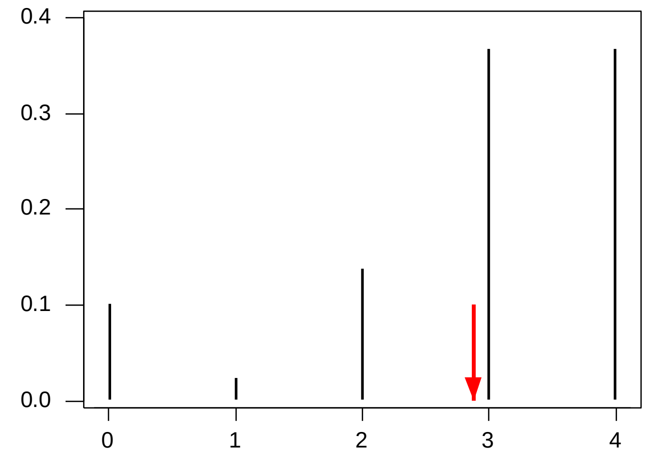




Exemple "1 Servidor i 4 connexions" amb VAD

En l'exemple del servidor i de la xarxa, analitzem l'esperança i variància de la variable *X* "número de connexions al servidor amb *n*=4"

k	$P_X(k)$	$k \cdot P_X(k)$	$(k-\mu)^2$	$(k-\mu)^2 \cdot P_X(k)$
0	0.10144	0	8.2944	0.84138
1	0.02304	0.02304	3.5344	0.08143
2	0.13824	0.27648	0.7744	0.10705
3	0.36864	1.10592	0.0144	0.00531
4	0.36864	1.47456	1.2544	0.46242
		$\mu=\sum k \cdot P_X(k)=2.88$		$\sigma^2=\sum(k-\mu)^2 \cdot P_X(k)=1.50$ $\sigma=\sqrt{\sigma^2}=1.22$



L'esperança 2.88 indica que s'espera un promig de 2.88 connexions si l'experiència es repetís un gran número de vegades. No s'ha d'arrodonir a un enter. L'esperança s'associa al centre de gravetat.

La desviació 1.22 informa sobre la magnitud de la dispersió de la variable. Major desviació suposa major probabilitat de buscar un valor allunyat del valor esperat. Menor desviació, major concentració.