Jeu de Passe-Dix

Paradoxe du jeu de Passe-Dix

À la fin du XVIème siècle, le Duc de Toscane aurait posé à Galilée le problème suivant au sujet du jeu de Passe-Dix. Ce jeu consiste à lancer trois dés et à observer la somme des points marqués.
« Au jeu de Passe-Dix, on a 6 manières possibles de faire le 11 comme le 12. En effet :
Somme = 11 : "6,4,1" "6,3,2" "5,5,1" "5,4,2 "5,3,3" "4,4,3"
Somme = 12 : "6,5,1" "6,4,2" "6,3,3" "5,5,2" "5,4,3" "4,4,4"
Pourtant, lorsqu'on joue un grand nombre de fois, on s'aperçoit que le 11 sort plus souvent que le 12. Qu'en pensez-vous ? »

Les paradoxes sont souvent à l'origine de nouvelles découvertes mathématiques. Il s'agit ici d'un paradoxe reposant sur les deux intuitions des probabilités, l'équiprobabilité pour des raisons de symétrie et l'approche dite fréquentiste : la fréquence de succès sur un grand nombre d'expériences est proche de la « probabilité » de succès.
Des enquêtes réalisées auprès d'élèves de collège semblent indiquer que l'intuition selon laquelle « la fréquence de succès se stabilise lorsque le nombre d'expériences augmente » n'est pas unanimement partagée.

Solution

L'erreur du Duc de Toscane est de penser que les différentes manières de faire le 11 comme le 12 ont la même probabilité de se produire.
Or, obtenir « 6,4,1 » est six fois plus probable qu'obtenir « 4,4,4 » et obtenir « 4,4,3 » est trois fois plus probable qu'obtenir « 4,4,4 ». En fait, on a équiprobabilité sur les 216 triplets possibles construits à partir de {1,…,6}. On a physiquement 3 dés, même s'ils sont indiscernables à l'œil, …

Plus formellement, afin de se ramener à l'équiprobabilité, on suppose que les trois dés sont « discernables » : A, B et C. Un résultat est alors le triplet des 3 nombres marqués par les dés A, B et C respectivement.

$\Omega = \{1,2,3,4,5,6\}^3 = \{(1,1,1), (1,1,2), \dots, (6,6,6)\} = [\![1,6]\!]^3$; on a 6^3 soit 216 résultats possibles équiprobables.

La probabilité d'avoir "6,4,1" est : $P(\{(6,4,1),(6,1,4),(1,6,4),(1,4,6),(4,6,1),(4,1,6)\}) = \frac{6}{216}$

De même pour "6,3,2" , "5,4,2" , "6,5,1" , "6,4,2" , "5,4,3" (trois nombres distincts).

La probabilité d'avoir "5,5,1" est alors : $P(\{(5,5,1), (5,1,5), (1,5,5)\}) = \frac{3}{216}$

De même pour "5,3,3" , "4,4,3" , "6,3,3" , "6,4,2" , "5,5,2" (deux nombres égaux).

La probabilité d'avoir “4,4,4” est alors : $\mathrm{P}\left(\left\{(4,4,4)\right\}\right)=\frac{1}{216}$.

On en déduit, en reprenant dans l'ordre les six manières possibles décrites dans la lettre du Duc de Toscane :

$$P(Somme=11)=\frac{6}{216}+\frac{6}{216}+\frac{3}{216}+\frac{6}{216}+\frac{3}{216}+\frac{3}{216}=\frac{27}{216}$$

$$P(Somme=12)=\frac{6}{216}+\frac{6}{216}+\frac{3}{216}+\frac{3}{216}+\frac{6}{216}+\frac{1}{216}=\frac{25}{216}$$

Finalement, la probabilité d'obtenir le « 11 » est supérieure à celle d'obtenir le « 12 » … mais il a fallu que le Duc de Toscane joue un grand nombre de fois pour observer une différence « significative » entre les fréquences d'obtention du « 11 » et du « 12 ».

Remarque sur la loi de probabilité de la v.a.r. donnant la somme des points marqués par les trois dés.
L'espace probabilisé que nous avons associé à l'expérience aléatoire est :
$\left(\Omega,\mathcal{P}(\Omega),P\right)$ où $\Omega=[\![1,6]\!]^3$ et P l'équiprobabilité sur $\left(\Omega,\mathcal{P}(\Omega)\right)$.
La v.a.r. *X*, somme des points marqués par les trois dés, est définie sur Ω et à valeurs dans $[\![3,18]\!]$;
la loi de probabilité P_X de X sur $[\![3,18]\!]$ est l'image de *P* par *X*, définie par :
$\forall k\in[\![3,18]\!],\ P_X\left(\{k\}\right)=P\left([X=k]\right)$, où l'on rappelle :
$[X=k]=X^{-1}\left(\{k\}\right)=\left\{(a,b,c)\in[\![1,6]\!]^3\,;\,a+b+c=k\right\}$.

Vers les simulations

On sait que, si l'on renouvelle l'expérience aléatoire *n* fois dans les mêmes conditions, la fréquence de sortie du 11, $F_{n,11}$, (resp. de sortie du 12, $F_{n,12}$,) est une v.a.r. définie sur l'espace produit $\left(\Omega,\mathcal{P}(\Omega),P\right)^{\otimes n}$, dont la loi est approximativement une loi normale de moyenne *p* et d'écart-type $\sqrt{\frac{p(1-p)}{n}}$, avec $p = 27/216 \simeq 0.125$ (resp. $p = 25/216 \simeq 0.116$) soit :

$$F_{n,11}\overset{approx}{\sim}\mathcal{N}\left(0.125,\frac{0.331}{\sqrt{n}}\right)\text{ et }F_{n,12}\overset{approx}{\sim}\mathcal{N}\left(0.116,\frac{0.320}{\sqrt{n}}\right).$$

On cherche la taille *n* de l'échantillon tel que $P\left(\left|F_{n,11}-\frac{27}{216}\right|\le\frac{1}{216}\right)\simeq 0.95$; on aura alors aussi $P\left(\left|F_{n,12}-\frac{25}{216}\right|\le\frac{1}{216}\right)\simeq 0.95$ et la fréquence d'obtenir le 11 sera supérieure à la fréquence d'observer le 12.
Comme on a $P\left(\left|F_{n,11}-0.125\right|\le 1.96\frac{0.331}{\sqrt{n}}\right)\simeq 0.95$, la taille de l'échantillon doit vérifier $1.96\frac{0.331}{\sqrt{n}}\simeq\frac{1}{216}$, soit $n\simeq 19600$.
En fait la probabilité d'observer une fréquence de sortie du 11 significativement supérieure à la fréquence de sortie du 12 au seuil de 5 % arrivera à partir d'un nombre d'expériences inférieur. En effet, on cherche la taille *n* de l'échantillon tel que $P\left(F_{n,11}-F_{n,12}>0\right)\simeq 0.95$.

Les v.a.r. $F_{n,11}$ et $F_{n,12}$ ne sont pas indépendantes (on a en effet $\sum_{k=3}^{18} F_{n,k} = 1$) mais la covariance entre les deux variables est faible (égale à $-\frac{25}{216} \times \frac{27}{216} \times \frac{1}{n}$), aussi on admet que la différence $D_n = F_{n,11} - F_{n,12}$ suit encore une loi normale.

Comme on a $E(D_n) = \frac{27}{216} - \frac{25}{216} = \frac{2}{216} \simeq 0.009$ et

$V(D_n) = V(F_{n,11}) - 2\operatorname{cov}(F_{n,11}, F_{n,12}) + V(F_{n,12}) \simeq \frac{(0.428)^2}{n}$, on en déduit :

$D_n \overset{approx}{\sim} N(0.009, \frac{0.428}{\sqrt{n}})$ et $U = \frac{D_n - 0.009}{0.428/\sqrt{n}} \overset{approx}{\sim} N(0,1)$.

D'une part, la lecture de la table de la loi normale centrée réduite permet d'écrire : $P(U > -1.645) \simeq 0.95$, d'autre part, on a

$$P(D_n > 0) = P\left(\frac{D_n - 0.009}{0.428/\sqrt{n}} > -\frac{0.009}{0.428/\sqrt{n}}\right) = P\left(U > -0.0216\sqrt{n}\right) \simeq 0.95,$$

la taille *n* de l'échantillon doit donc vérifier $-0.0216\sqrt{n} \simeq -1.645$ d'où $n \simeq 5780$.

A partir d'environ 5800 expériences, la probabilité que la fréquence de sortie du 11 soit supérieure à la fréquence de sortie du 12 est approximativement égale à 0.95.

Simulations

On souhaite, par simulations, approcher la probabilité que X soit égal à 11, que X soit égal à 12. Voici, sur Voyage 200, deux programmes (des fonctions en fait) donnant, pour un nombre de simulations n fixé, le nombre de fois où 11 et le nombre de fois où 12 sont apparus.

Commentaires :
- une fois un programme écrit, il est exécuté en affectant un entier à n (passdix(100) ou passe(100) si l'on veut réaliser 100 simulations)
- il est possible de cumuler les résultats de différentes simulations pour avoir un résultat sur un nombre de simulations plus important
- dans chacun des deux programmes, i désigne la i^ème simulation (de 1 à n), j (resp. k) est le "compteur" de sorties du 11 (resp. 12),
- x est la valeur simulée de la variable aléatoire X dans le premier programme, x est la liste des n valeurs simulées de la variable aléatoire X dans le deuxième

passdix(n)
Func
Local i, j, k, x
0→j
0→k
For i, 1, n
rand(6)+rand(6)+rand(6) → x
If x = 11 Then
j+1→j
Else
If x = 12 Then
k+1→k
EndIf
EndIf
EndFor
{j,k}
EndFunc
	passe(n)
	Func
	Local i, j, k, x
	seq(rand(6)+ rand(6)+ rand(6), i, 1, n)→x
	0 → j
	0 → k
	For i, 1, n
	If x[i] = 11 Then
	j + 1 → j
	Else
	If x[i] = 12 Then
	k + 1 → k
	EndIf
	EnfIf
	EndFor
	{j, k}
	EndFunc