

## Capes 2016, épreuve 2, problème 2 ou « QPUC en ligne »

En 2013, j'avais mis sur mon site un problème inspiré du jeu « Question pour un Champion en ligne », sujet que le lecteur, s'il le souhaite, pourra retrouver en cherchant un peu. Je reproduis ci-dessous l'introduction de ce problème :

« QPUC « en ligne » se déroule ainsi :

Six internautes disputent une première manche dite « 9 points gagnants ». Les quatre premiers sont qualifiés pour une deuxième manche dite « 4 à la suite ». A l'issue de cette deuxième manche, deux finalistes sont retenus pour s'affronter dans une troisième manche « Face à face ».

Les deux finalistes du « Face à face » doivent répondre à un même questionnaire, chacun ayant la main à tour de rôle pendant quelques secondes. Ce questionnaire est tiré au hasard parmi une liste de questionnaires préétablie et en nombre relativement limitée. Aussi, il arrive aux internautes assidus de retrouver au cours d'une finale un questionnaire auquel ils ont déjà répondu.

C'est pourquoi des petits malins prennent soin d'archiver à l'issue de chacun de leurs « Face à face » les questionnaires auxquels ils se mesurent. Lorsque ces petits malins ont la chance de retrouver une série déjà connue, ils désintègrent leur adversaire en trois coups de cuillers à pot puisqu'ils connaissent d'avance toutes les réponses. Il est alors frustrant pour l'adversaire malchanceux de voir la réponse inscrite par son rival alors que la question n'est même pas entièrement posée (à moins qu'il ne soit lui-même un petit malin chanceux, auquel cas c'est le premier à reconnaître le quiz qui gagne ... ) ».

### Le sujet

(Non reproduit ici). Voir par exemple le site du jury du CAPES <http://capes-math.org/> ...

On reconnaîtra que la situation étudiée dans le problème « épreuve 2, problème 2 » du CAPES 2016 est la même que celle décrite ci-dessus. Les « boules blanches » représentent les questionnaires déjà connus et archivés. Les « boules rouges » représentent les nouveaux questionnaires encore non archivés. Lorsqu'un tel questionnaire est tiré, il est à son tour archivé (la « boule rouge » devient blanche).

Il n'y a aucun inconvénient dans le problème CAPES 2016 à supposer que  $b=0$  ;  $r=N$  (c'est-à-dire qu'au début de l'expérience, toutes les boules sont rouges). Certaines formules y gagnent en limpidité.

Cet intéressant sujet du CAPES 2016 va nettement plus loin que mon problème de 2013 et obtient des résultats incontestablement plus gratifiants.

## Éléments de correction

## Partie A : Un cas particulier

## I. et II. Deux expériences aléatoires

Les modélisations (non reproduites ici) à l'aide d'arbres pondérés amènent aux résultats suivants :

Valeurs de $X$	1	2	3	4		Valeurs de $Y$	0	1	2	3
Probabilité	$\frac{2}{5} = \frac{50}{125}$	$\frac{9}{25} = \frac{45}{125}$	$\frac{24}{125}$	$\frac{6}{125}$		Probabilité	$\frac{8}{125}$	$\frac{57}{125}$	$\frac{54}{125}$	$\frac{6}{125}$

On obtient :  $E(X)_{\text{gijulia2016}} = \frac{236}{125} = 1,89$  à  $10^{-2}$  près pour la première expérience aléatoire et  $E(Y) = \frac{183}{125} = 1,46$  à  $10^{-2}$  près pour la deuxième.

Dans chaque cas, l'espérance peut être interprétée comme la limite de la moyenne des résultats au cours de séries de  $N$  expériences, quand le nombre  $N$  d'expériences augmente indéfiniment.

Ce programme simule la deuxième expérience. Les variables  $b$  et  $r$  désignent les nombres de boules blanches et rouges respectivement.

La variable  $x$  représente le nombre de tirages effectués pour obtenir une première boule blanche.

```

* cp2016a 1/9
Define cp2016a(b,r)=
Prgm
Local x
1 → x
While randInt(1,b+r) ≤ r
1 + x → x
©gilbertjulia2016
r - 1 → r
b + 1 → b
EndWhile
Disp x
EndPrgm

```

Ce programme simule la deuxième expérience. Les variables  $b$  et  $r$  désignent les nombres de boules blanches et rouges respectivement. La variable  $n$  désigne le nombre de tirages. Ici, on simule des séries de trois tirages.

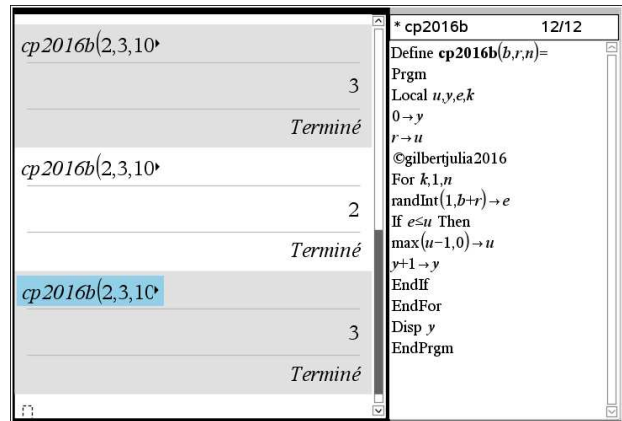
La variable  $y$  représente le nombre de boules rouges obtenues.

```

* cp2016b 12/12
Define cp2016b(b,r,n)=
Prgm
Local u,y,e,k
0 → y
r → u
©gilbertjulia2016
For k,1,n
randInt(1,b+r) → e
If e ≤ u Then
max(u-1,0) → u
y + 1 → y
EndIf
EndFor
Disp y
EndPrgm

```

Ici on simule au lieu de séries de trois tirages des séries de 10 tirages. Il est évidemment plus fréquent d'obtenir «  $Y = 3$  »



### Partie B : généralisation, première expérience aléatoire

**I. 1.** Le nombre de tirages effectués avant de tirer une boule blanche est nécessairement inférieur ou égal à  $(r + 1)$ . Les valeurs prises par  $X$  sont les  $(r + 1)$  valeurs de l'ensemble  $\{1, 2, \dots, r + 1\}$ .

**2.** Pour  $1 \leq k \leq r$  :  $(X = k) = A_1 \cap \dots \cap A_{k-1} \cap \overline{A_k}$ .

Se distingue des autres et fait exception un évènement :  $(X = r + 1) = A_1 \cap \dots \cap A_r$

**3.1.**  $B$  étant un évènement de probabilité non nulle, la probabilité conditionnelle de  $A$  sachant  $B$  est le quotient  $P_B(A) = \frac{P(A \cap B)}{P(B)}$

**3.2.** Pour tout  $j$  tel que  $j \leq i - 1$ ,  $(B_1 \cap \dots \cap B_j) \supset (B_1 \cap \dots \cap B_{i-1})$ . Par conséquent :  $P(B_1 \cap \dots \cap B_j) \geq P(B_1 \cap \dots \cap B_{i-1})$ .

Si  $P(B_1 \cap \dots \cap B_{i-1}) > 0$ , alors *a fortiori* pour toute intersection  $B_1 \cap \dots \cap B_j$  d'un plus petit nombre d'évènements  $P(B_1 \cap \dots \cap B_j) > 0$  ce qui justifie l'existence de probabilités conditionnelles selon ces évènements.

Notamment :  $P(B_1 \cap \dots \cap B_{i-1} \cap B_i) \stackrel{g\text{julia}2016}{=} P(B_1 \cap \dots \cap B_{i-1}) \times P_{B_1 \cap \dots \cap B_{i-1}}(B_i)$  puis :

$$P(B_1 \cap \dots \cap B_{i-2} \cap B_{i-1}) = P(B_1 \cap \dots \cap B_{i-2}) \times P_{B_1 \cap \dots \cap B_{i-2}}(B_{i-1}).$$

$$P(B_1 \cap \dots \cap B_{i-2}) = \dots$$

...

$$P(B_1 \cap B_2 \cap B_3) = P(B_1 \cap B_2) \times P_{B_1 \cap B_2}(B_3)$$

$$P(B_1 \cap B_2) = P(B_1) \times P_{B_1}(B_2)$$

On obtient par multiplication puis simplification des termes deux fois apparus :

$$P(B_1 \cap \dots \cap B_{i-1} \cap B_i) \stackrel{g\text{julia}2016}{=} P_{B_1 \cap \dots \cap B_{i-1}}(B_i) \times P_{B_1 \cap \dots \cap B_{i-2}}(B_{i-1}) \times \dots \times P_{B_1}(B_2) \times P(B_1)$$

**4.1.** Si l'urne contient  $r$  boules rouges, l'évènement :  $A_1 \cap A_2 \cap \dots \cap A_r$  est un évènement de probabilité non nulle, ce qui justifie l'existence de probabilités conditionnelles concernant des intersections de ces  $A_j$ .

Selon le protocole établi, pour tout entier  $j$  tel que  $1 \leq j \leq r$  :  $P_{A_1 \cap \dots \cap A_j}(A_{j+1}) = \frac{r-j}{N}$  puisqu'on sait que  $j$  boules rouges ont déjà été remplacées.

On obtient :  $P(X=1) = P(\overline{A_1}) = 1 - \frac{r}{N}$  puis :

$$P(X=2) = P(A_1 \cap \overline{A_2}) = P(A_1) \times P_{A_1}(\overline{A_2}) = P(A_1) \times (1 - P_{A_1}(A_2)) = \frac{r}{N} \times \left(1 - \frac{r-1}{N}\right)$$

De façon plus générale, pour  $1 \leq k \leq r$  :

$$P(X=k) = P(A_1 \cap \dots \cap A_{k-1} \cap \overline{A_k}) = P(A_1) \times P_{A_1}(A_2) \times \dots \times P_{A_1 \cap \dots \cap A_{k-2}}(A_{k-1}) \times (1 - P_{A_1 \cap \dots \cap A_{k-1}}(A_k)).$$

Ce qui donne :  $P(X=k) = \frac{r}{N} \times \frac{r-1}{N} \times \dots \times \frac{r-k+1}{N} \times \left(1 - \frac{r-k}{N}\right)$ .

Soit aussi :  $P(X=k) = \frac{\prod_{j=0}^{k-1} (r-j)}{N^{k-1}} - \frac{\prod_{j=0}^k (r-j)}{N^k}$ , expressions équivalentes aux expressions avec factorielles données dans l'énoncé.

On note au passage que si on pose pour  $0 \leq k \leq r$  :  $p_k = \frac{r!}{(r-k)!N^k}$  on obtient :  $P(X=k) = p_{k-1} - p_k$

Quant à l'évènement exceptionnel :

$$P(X=r+1) = P(A_1 \cap \dots \cap A_{r-1} \cap A_r) = P(A_1) \times P_{A_1}(A_2) \times \dots \times P_{A_1 \cap \dots \cap A_{r-1}}(A_r)$$

Soit :  $P(X=r+1) = P(A_1 \cap \dots \cap A_{r-1} \cap A_r) = \frac{r}{N} \times \frac{r-1}{N} \times \dots \times \frac{1}{N} = \frac{r!}{N^r}$

$$5.1. \sum_{k=1}^n k(p_{k-1} - p_k) = \sum_{k=1}^n k p_{k-1} - \sum_{k=1}^n k p_k.$$

En changeant dans la première somme l'indexation par :  $j = k-1$ , on obtient

$$\sum_{k=1}^n k(p_{k-1} - p_k) = \sum_{j=0}^{n-1} (j+1)p_j - \sum_{k=1}^n k p_k \text{ soit : } \sum_{k=1}^n k(p_{k-1} - p_k) = \left( \sum_{j=0}^{n-1} p_j + \sum_{j=1}^{n-1} j p_j \right) - \sum_{k=1}^n k p_k.$$

On reconnaît deux sommes identiques à un index près, l'index  $n$  qui figure dans la dernière somme et non dans la deuxième. Il reste :  $\sum_{k=1}^n k(p_{k-1} - p_k) = \sum_{j=0}^{n-1} p_j - n p_n$

5.2. L'espérance mathématique de  $X$  est :  $E(X) = \sum_{k=1}^{k=r+1} k P(X=k)$  qu'on peut exprimer avec les notations des

questions précédentes par :  $E(X) = \sum_{k=1}^{k=r} k(p_{k-1} - p_k) + (r+1)p_r$ .

$$\text{Compte tenu de 5.1 : } E(X) = \left( \sum_{k=0}^{k=r-1} p_k \right) - r p_r + (r+1)p_r = \sum_{k=0}^{k=r} p_k$$

Or on a posé :  $p_k = \frac{r!}{(r-k)!N^k} = \binom{r}{k} \frac{k!}{N^k}$  ce qui achève la preuve.

Il est possible de contrôler la concordance des résultats obtenus dans le cas particulier avec ceux issus de l'étude du cas général.

Define  $p(k) = \frac{(3-(k-1))!}{5^{k-1}} - \frac{(3-k)!}{5^k}$

$\{p(1), p(2), p(3)\}$   $\left\{ \frac{2}{5}, \frac{9}{25}, \frac{24}{125} \right\}$

$\frac{3!}{5^3}$   $\frac{6}{125}$

$\sum_{k=1}^3 (k \cdot p(k)) + \frac{4 \cdot 6}{125}$   $\frac{236}{125}$

©gilbertjulia2016

## Seconde expérience aléatoire

**II. 1.1.** Il est impossible d'obtenir plus de boules rouges que de boules tirées.

**1.2.** Il est d'autre part impossible de tirer plus de boules rouges qu'il n'y en a initialement dans l'urne.

**1.3.**  $P(Y_n = 0) = \frac{(N-r)^n}{N^n}$ .

**2.** La probabilité conditionnelle dont il est question a un sens du moment que  $i \leq n-1$  et que  $i \leq r$ . Dans le cas contraire, l'évènement  $Y_{n-1} = i$  a une probabilité nulle. Cette réserve étant faite :

$$P_{Y_{n-1}=i}(Y_n = k) = \begin{cases} \frac{r-i}{N} & \text{si } k = i+1 \\ 1 - \frac{r-i}{N} = \frac{N-r+i}{N} & \text{si } k = i \text{ puisqu'au } n\text{-ème tirage, on peut ou non obtenir une boule} \\ 0 & \text{sinon} \end{cases}$$

rouge de plus qu'on en a déjà.

Pour  $1 \leq k \leq n$ , l'évènement  $Y_n = k$  est une réunion de deux évènements disjoints :

$$(Y_n = k) = ((Y_{n-1} = k) \cap (Y_n = k)) \cup ((Y_{n-1} = k-1) \cap (Y_n = k))$$

Cependant, l'un des deux ou les deux évènements en question peuvent être impossibles.

Par exemple,  $(Y_n = n) = ((Y_{n-1} = n-1) \cap (Y_n = n))$  car  $(Y_{n-1} = n)$  est impossible

Sous réserve d'existence ( $k \leq n-1$  et  $k \leq r$ ) des probabilités conditionnelles dont on parle :

$$P(Y_n = k) = P(Y_{n-1} = k) \times P_{Y_{n-1}=k}(Y_n = k) + P(Y_{n-1} = k-1) \times P_{Y_{n-1}=k-1}(Y_n = k)$$

$$P(Y_n = k) = P(Y_{n-1} = k) \times \frac{N-r+k}{N} + P(Y_{n-1} = k-1) \times \frac{r-k+1}{N}$$

$$P(Y_n = k) =_{\text{gilbia2016}} P(Y_{n-1} = k) \times \frac{N-r+k}{N} + P(Y_{n-1} = k-1) \times \frac{r-k+1}{N}$$

Ces formules s'étendent aux cas où cet usage des probabilités conditionnelles utilisées pour les établir n'est plus légitime parce que les probabilités correspondantes sont nulles.

Par exemple :  $P(Y_n = n) = P(Y_{n-1} = n-1) \times \frac{r+1-n}{N}$  est compatible avec la formule générale car

$$P(Y_{n-1} = n) = 0$$

$$3. E(Y_n) = \sum_{k=1}^n k P(Y_n = k) = \sum_{k=1}^n k P(Y_{n-1} = k) \times \frac{N-r+k}{N} + \sum_{k=1}^n k P(Y_{n-1} = k-1) \times \frac{r-k+1}{N}$$

D'une part :  $\sum_{k=1}^n k P(Y_{n-1} = k-1) \times \frac{r-k+1}{N} = \sum_{j=0}^{n-1} (j+1) P(Y_{n-1} = j) \times \frac{r-j}{N}$  en posant  $j = k-1$

Soit  $\sum_{k=1}^n k P(Y_{n-1} = k-1) \times \frac{r-k+1}{N} = \sum_{k=0}^{n-1} (k+1) P(Y_{n-1} = k) \times \frac{r-k}{N}$  avec la notation  $k$  au lieu de  $j$ .

D'autre part :  $\sum_{k=1}^n k P(Y_{n-1} = k) \times \frac{N-r+k}{N} = \sum_{k=0}^{n-1} k P(Y_{n-1} = k) \times \frac{N-r+k}{N}$ , car le terme ajouté (d'index zéro) est nul et le terme enlevé (d'index  $n$ ) aussi.

$$E(Y_n) = \sum_{k=0}^{n-1} k P(Y_{n-1} = k) \times \frac{N-r+k}{N} + \sum_{k=0}^{n-1} (k+1) P(Y_{n-1} = k) \times \frac{r-k}{N}$$

$$E(Y_n) = \sum_{k=0}^{n-1} P(Y_{n-1} = k) \times \frac{k(N-r+k) + (k+1)(r-k)}{N}$$

$$E(Y_n) = \sum_{k=0}^{n-1} P(Y_{n-1} = k) \times \frac{k(N-1) + r}{N}$$

Enfinement :  $E(Y_n) = \frac{r}{N} \left( \sum_{k=0}^{n-1} P(Y_{n-1} = k) \right) + \frac{N-1}{N} \left( \sum_{k=0}^{n-1} k P(Y_{n-1} = k) \right) = \frac{r}{N} + \frac{N-1}{N} E(Y_{n-1})$

**4.1.** La variable aléatoire  $Y_1$  prend les valeurs 0 et 1 avec les probabilités  $\frac{N-r}{N}$  et  $\frac{r}{N}$  respectivement. De ce

fait :  $E(Y_1) = \frac{r}{N} = r \left( 1 - \frac{N-1}{N} \right)$ . La formule proposée est exacte au rang 1. (Elle l'est déjà au rang zéro si on se rappelle que  $Y_0$  est conventionnellement la variable nulle).

Supposons-la exacte au rang  $n-1$ .

Alors :  $E(Y_n) = \frac{r}{N} + \frac{N-1}{N} \left( r - r \left( \frac{N-1}{N} \right)^{n-1} \right) = \frac{r + r(N-1)}{N} - \frac{N-1}{N} \left( \frac{N-1}{N} \right)^{n-1} = r \left( 1 - \left( \frac{N-1}{N} \right)^n \right)$

Elle est encore exacte au rang  $n$ . Elle est donc exacte pour tout entier naturel  $n$ .

Il est possible de contrôler la concordance des résultats obtenus dans le cas particulier avec ceux issus de l'étude du cas général.

	( 5 25 125 )
$\frac{3!}{5^3}$	$\frac{6}{125}$
$\sum_{k=1}^3 (k \cdot p(k)) + \frac{4 \cdot 6}{125}$	$\frac{236}{125}$
©gilbertjulia2016	
$3 \cdot \left( 1 - \left( 1 - \frac{1}{5} \right)^3 \right)$	$\frac{183}{125}$
1	

La suite  $n \mapsto \left(\frac{N-1}{N}\right)^n$  est une suite géométrique de raison plus petite que 1, elle converge vers zéro.

Donc  $\lim_{n \rightarrow +\infty} E(Y_n) = r$

4.2.  $r - E(Y_n) = r \left(\frac{N-1}{N}\right)^n$ . On obtient :  $r - E(Y_n) = r \left(\frac{N-1}{N}\right)^n < \frac{1}{4} \Leftrightarrow \ln r + n(\ln(N-1) - \ln N) < -2\ln 2$

La condition :  $n > \frac{2\ln 2 + \ln r}{\ln N - \ln(N-1)}$  assure que  $r - E(Y_n) < \frac{1}{4}$ .

On peut choisir pour  $n_0$  le premier entier supérieur au nombre  $\frac{2\ln 2 + \ln r}{\ln N - \ln(N-1)}$ .

Par exemple, pour  $b = 2 ; r = 3, n_0 = 12$ . Pour  $b = 20 ; r = 30, n_0 = 237$

III.1.  $k$  étant un entier strictement positif donné,  $Y_k$  prend les valeurs  $0, 1, \dots, k$ . Les évènements  $(Y_k = i)$ , pour  $0 \leq i \leq k$ , deux à deux disjoints et de réunion certaine, forment un système complet d'évènements.

$$A_{k+1} = \bigcup_{i=0, \dots, k} (A_{k+1} \cap (Y_k = i)) \text{ et } P(A_{k+1}) = \sum_{i=0}^k P(A_{k+1} \cap (Y_k = i))$$

Tant que  $i \leq r$  :  $P(A_{k+1} \cap (Y_k = i)) = P_{Y_k=i}(A_{k+1}) = \frac{r-i}{N} P(Y_k = i)$  et lorsque  $i > r$   $P(Y_k = i) = 0$

Que  $k \leq r$  ou non, il est donc légitime d'écrire :  $P(A_{k+1}) = \sum_{i=0}^k P(A_{k+1} \cap (Y_k = i)) = \sum_{i=0}^k \frac{r-i}{N} P(Y_k = i)$

III.2. On obtient ainsi :  $P(A_{k+1}) = \frac{r}{N} \sum_{i=0}^k P(Y_k = i) - \frac{1}{N} \sum_{i=0}^k i P(Y_k = i) = \frac{r}{N} - \frac{1}{N} E(Y_k) = \frac{r}{N} \left(1 - \frac{1}{N}\right)^k$   
gjuia2016

IV.1. Si on considère la variable aléatoire  $Y_n(Y_n - 1)$ , elle est image de  $Y_n$  par la fonction  $x \mapsto f(x) = x(x-1)$ .

D'après le théorème du transfert :  $E(Y_n(Y_n - 1)) = \sum_{k=0}^n f(k)P(Y_n = k) = \sum_{k=2}^n k(k-1)P(Y_n = k)$  (les termes de rangs 0 et 1 ont une contribution nulle)

$$\sum_{k=2}^n k(k-1)P(Y_n = k) = \sum_{k=2}^n k(k-1)P(Y_{n-1} = k) \times \frac{N-r+k}{N} + \sum_{k=2}^n k(k-1)P(Y_{n-1} = k-1) \times \frac{r-k+1}{N}$$

D'une part :  $\sum_{k=2}^n k(k-1)P(Y_{n-1} = k-1) \times \frac{r-k+1}{N} = \sum_{j=1}^{n-1} ((j+1)j)P(Y_{n-1} = j) \times \frac{r-j}{N}$  en posant  $j = k-1$   
gj2016

Cette somme s'écrit  $\sum_{k=1}^{n-1} ((k+1)k)P(Y_{n-1} = k) \times \frac{r-k}{N}$  avec la notation  $k$ .

D'autre part :  $\sum_{k=2}^n k(k-1)P(Y_{n-1} = k) \times \frac{N-r+k}{N} = \sum_{k=1}^{n-1} k(k-1)P(Y_{n-1} = k) \times \frac{N-r+k}{N}$ , car le terme ajouté (d'index 1) est nul et le terme enlevé (d'index  $n$ ) aussi.

$$E(Y_n(Y_n - 1)) = \sum_{k=0}^{n-1} k(k-1)P(Y_{n-1} = k) \times \frac{N-r+k}{N} + \sum_{k=0}^{n-1} (k+1)kP(Y_{n-1} = k) \times \frac{r-k}{N}$$

$$E(Y_n(Y_n - 1)) = \sum_{k=0}^{n-1} P(Y_{n-1} = k) \times \frac{k(k-1)(N-r+k) + (k+1)k(r-k)}{N}$$

Or  $(k-1)(N-r+k) + (k+1)(r-k) = 2r - N + k(N-2) = (k-1)(N-2) + 2r - 2$ . La formule de sommation obtenue est équivalente à celle de l'énoncé.

Ainsi :

$$E(Y_n(Y_n - 1)) = \frac{N-2}{N} \sum_{k=0}^{n-1} k(k-1)P(Y_{n-1} = k) + \frac{2(r-1)}{N} \sum_{k=0}^{n-1} kP(Y_{n-1} = k) = \frac{N-2}{N} E(Y_{n-1}(Y_{n-1} - 1)) + \frac{2(r-1)}{N} E(Y_{n-1})$$

On en déduit la formule de récurrence de l'énoncé, vu l'expression déjà obtenue de  $E(Y_{n-1})$

**IV.3.** Au rang 1,  $E(Y_1(Y_1 - 1)) = 0$  conformément à la formule qui se trouve de ce fait initialisée. (Elle l'est déjà au rang zéro, compte tenu de la convention « variable aléatoire nulle »).

Supposons la formule exacte au rang  $n-1$ . C'est-à-dire, supposons :

$$E(Y_{n-1}(Y_{n-1} - 1)) = r(r-1) \left( 1 + \left( \frac{N-2}{N} \right)^{n-1} - 2 \left( \frac{N-1}{N} \right)^{n-1} \right). \text{ Alors :}$$

$$E(Y_n(Y_n - 1)) = \frac{N-2}{N} \left( r(r-1) \left( 1 + \left( \frac{N-2}{N} \right)^{n-1} - 2 \left( \frac{N-1}{N} \right)^{n-1} \right) \right) + \frac{2r(r-1)}{N} \left( 1 - \left( \frac{N-1}{N} \right)^{n-1} \right)$$

$$E(Y_n(Y_n - 1)) = r(r-1) \left( \left( \frac{N-2}{N} + \left( \frac{N-2}{N} \right)^n - \frac{2(N-2)}{N} \left( \frac{N-1}{N} \right)^{n-1} \right) + \frac{2}{N} \left( 1 - \left( \frac{N-1}{N} \right)^{n-1} \right) \right)$$

$$\text{soit : } E(Y_n(Y_n - 1)) = r(r-1) \left( 1 + \left( \frac{N-2}{N} \right)^n - \frac{2(N-1)}{N} \left( \frac{N-1}{N} \right)^{n-1} \right) = r(r-1) \left( 1 + \left( \frac{N-2}{N} \right)^n - 2 \left( \frac{N-1}{N} \right)^n \right).$$

La formule est encore exacte au rang suivant  $n$ .

Elle l'est donc pour tout entier naturel  $n$ .

**IV.4.** D'après la formule de Koenig :  $V(Y_n) = E(Y_n^2) - (E(Y_n))^2$ . Ce qui donne aussi par linéarité de

l'espérance :  $V(Y_n) = E(Y_n(Y_n - 1)) + E(Y_n) - (E(Y_n))^2$

$$V(Y_n) = r(r-1) \left( 1 + \left( \frac{N-2}{N} \right)^n - 2 \left( \frac{N-1}{N} \right)^n \right) + r \left( 1 - \left( \frac{N-1}{N} \right)^n \right) - r^2 \left( 1 - \left( \frac{N-1}{N} \right)^n \right)^2$$



**V.1.** Dans l'expression de la variance, les suites géométriques de raisons respectives  $\frac{N-1}{N}$  et  $\frac{N-2}{N}$ , raisons toutes deux positives et plus petites que 1, convergent vers zéro. Donc  $\lim_{n \rightarrow +\infty} V(Y_n) = 0$

**V.2.** D'après l'inégalité de Bienaymé Tchebychev, pour tout réel  $\alpha$  :  $P(|Y_n - E(Y_n)| \geq \alpha) \leq \frac{V(Y_n)}{\alpha^2}$ . La suite à termes positifs  $(P(|Y_n - E(Y_n)| \geq \alpha))_{n \in \mathbb{N}}$  est majorée par une suite de limite nulle. Elle converge elle-même vers zéro.

**V.3.** Puisque l'urne contient  $r$  boules rouges, quelle que soit la valeur de  $n$ , les valeurs prises par  $Y_n$  appartiennent à l'ensemble  $\{0, 1, \dots, r\}$  (ce sont toutes ces valeurs si  $n \geq r$ ).

Ainsi quel que soit l'entier  $n$ ,  $0 \leq Y_n \leq r$  et  $0 \leq E(Y_n) \leq r$ .

Par conséquent :  $|Y_n - r| = r - Y_n$  et  $|r - E(Y_n)| = r - E(Y_n)$

L'évènement «  $|Y_n - r| - |r - E(Y_n)| \geq \alpha$  » est le même que l'évènement «  $E(Y_n) - Y_n \geq \alpha$  », évènement inclus dans «  $|E(Y_n) - Y_n| \geq \alpha$  ». Puisque d'après **V.2.** la probabilité de ce dernier évènement tend vers zéro quand  $n$  tend vers l'infini, il en est de même de la probabilité de «  $|Y_n - r| - |r - E(Y_n)| \geq \alpha$  ».

**V.4.** On a vu que les valeurs prises par  $Y_n$  appartiennent à l'ensemble  $\{0, 1, \dots, r\}$ .

Pour tout entier  $n \geq n_0$  :  $E(Y_n) \geq r - \frac{1}{4}$ .

Pour de tels entiers, de deux choses l'une, ou bien  $Y_n = r$  et dans ce cas  $(r - Y_n) - (r - E(Y_n)) \leq \frac{1}{4}$ , ou bien

$Y_n \leq r - 1$  et dans ce cas  $(r - Y_n) - (r - E(Y_n)) \geq \frac{3}{4}$

Si on choisit  $\alpha = \frac{1}{2}$  : l'inégalité  $(r - Y_n) - (r - E(Y_n)) \geq \frac{1}{2}$  a lieu si et seulement si  $Y_n$  prend une valeur de  $\{0, 1, \dots, r - 1\}$  (c'est-à-dire une valeur différente de  $r$ ).

De ce fait :  $P(E(Y_n) - Y_n) \geq \frac{1}{2} = P(Y_n \neq r)$

**5.** On peut conclure que  $\lim_{n \rightarrow \infty} P(Y_n \neq r) = 0$ . En augmentant indéfiniment la valeur de  $n$ , il devient presque sûr que à chaque expérience toutes les boules rouges sont extraites de l'urne.

Illustration à l'aide d'une simulation

Le programme **cp2016b** a été modifié de façon à simuler une série de **e** expériences identiques à la deuxième expérience aléatoire.

Ci-contre, on a choisi  $b = 20 ; r = 30 ; n = 100$ .

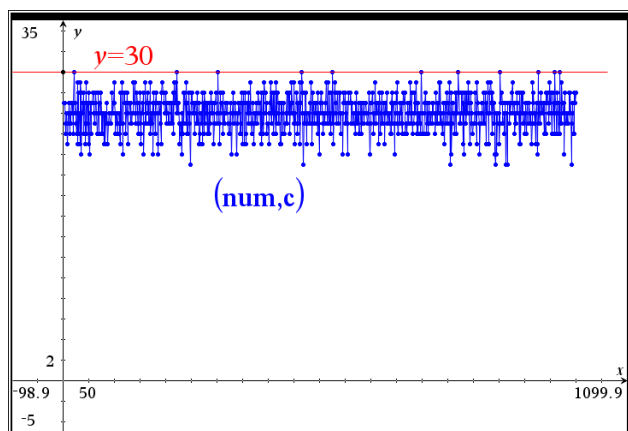
L'urne contient 30 boules rouges et 20 boules blanches. La liste **c** enregistre les résultats portant sur 1000 expériences identiques.

On a calculé la moyenne des résultats et on l'a comparée avec l'espérance théorique.

<code>cp2016b(20,30,100,1000)</code>	Terminé	"cp2016b" enregistré, effectué
<code>c</code>	<code>{ 27,27,27,24,25,26,23,23,26,25,27,24,26,25,26,27</code>	Define <b>cp2016b</b> ( <i>b,r,n,e</i> )= Prgm Local <i>u,y,e,k,j,m</i> newList( <i>e</i> )→ <i>c</i> ©gilbertjulia2016 For <i>j</i> ,1, <i>e</i> 0→ <i>y</i> <i>r</i> → <i>u</i> For <i>k</i> ,1, <i>n</i> randInt(1, <i>b+r</i> )→ <i>m</i> If <i>m</i> ≤ <i>u</i> Then max( <i>u</i> -1,0)→ <i>u</i> <i>y</i> +1→ <i>y</i> EndIf EndFor <i>y</i> → <i>c</i> [ <i>j</i> ] EndFor EndPrgm
$\frac{\text{sum}(c)}{1000}$	26.006	
$30 \cdot \left(1 - \left(\frac{49}{50}\right)^{100}\right)$	26.0214	

©gilbertjulia2016

Une représentation graphique de la liste **c** illustre assez bien ce qu'il se passe. On obtient un nombre de boules rouges qui grosso modo varie entre 20 et 30 ; il est rare d'obtenir les 30 boules rouges, mais parfois cela arrive.



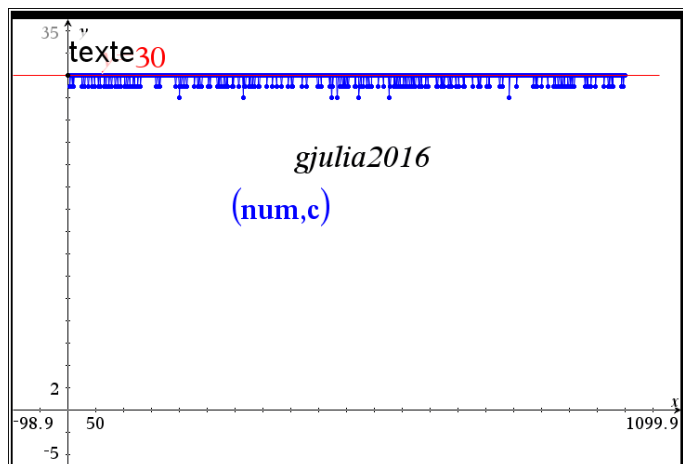
On a relancé le programme avec  $n = 250$  (et toujours 1000 expériences identiques).

Cet entier est choisi car il vérifie la condition  $n \geq n_0 = 237$

On constate que la moyenne des résultats est devenue plus voisine de 30. Sur cette simulation, l'écart avec 30 est  $< 0,25$ .

<code>cp2016b(20,30,100,1000)</code>	Terminé	cp2016b	0/15
<code>c</code>	<code>{ 27,27,27,24,25,26,23,23,26,25,27,26,27,24,27</code>	Define <b>cp2016b</b> ( <i>b,r,n,e</i> )= Prgm Local <i>u,y,e,k,j,m</i> newList( <i>e</i> )→ <i>c</i> ©gilbertjulia2016 For <i>j</i> ,1, <i>e</i> 0→ <i>y</i> <i>r</i> → <i>u</i> For <i>k</i> ,1, <i>n</i> randInt(1, <i>b+r</i> )→ <i>m</i> If <i>m</i> ≤ <i>u</i> Then max( <i>u</i> -1,0)→ <i>u</i> <i>y</i> +1→ <i>y</i> EndIf EndFor <i>y</i> → <i>c</i> [ <i>j</i> ] EndFor EndPrgm	
$\frac{\text{sum}(c)}{1000}$	26.006		
$30 \cdot \left(1 - \left(\frac{49}{50}\right)^{100}\right)$	26.0214		
©gilbertjulia2016			
<code>cp2016b(20,30,250,1000)</code>	Terminé		
$\frac{\text{sum}(c)}{1000}$	29.803		
$30 \cdot \left(1 - \left(\frac{49}{50}\right)^{250}\right)$	29.8079		

Le résultat 30 est devenu fréquent, les résultats inférieurs à 30 (en l'occurrence on observe 28 et 29) se sont raréfiés.



Maintenant,  $n = 500$ . Sur cette simulation, on ne remarque que trois essais au cours desquels il manque une seule boule rouge. Dans tous les autres cas, les 30 boules rouges ont été obtenues.

Si on augmentait encore la valeur de  $n$ , l'évènement «  $Y_n \neq 30$  » deviendrait rarissime ...

