On considère l'ensemble des familles de deux enfants. On dispose des informations suivantes :

- Dans la moitié d'entre elles, les enfants sont du même sexe.
- Dans 35% d'entre elles, l'enfant le plus jeune est un garçon.
- Dans 55% des cas, l'enfant le plus âgé est une fille.

Quelle est la proportion de garçons dans ces familles de 2 enfants ?

Analyse

L'exercice peut s'avérer délicat en ce sens que l'on doit bien distinguer les raisonnements portant sur les familles de ceux portant sur les enfants.
Nous fournissons deux approches. La seconde, bien que plus « probabiliste » que la première, est cependant moins détaillée, l'essentiel des difficultés se trouvant abordé dans la première approche.

Résolution

1ère approche

Dans les notations suivantes, l'indice 1 correspond à l'enfant le plus jeune et l'indice 2 à l'enfant le plus âgé.

Notons :

- N le nombre total de familles de 2 enfants.
- $N(G_1,G_2)$ le nombre de familles dont l'enfant le plus jeune et l'enfant le plus âgé sont des garçons.
- $N(G_1,F_2)$ le nombre total de familles dont l'enfant le plus jeune est un garçon et dont l'enfant le plus âgé est une fille.
- $N(F_1,G_2)$ le nombre total de familles dont l'enfant le plus jeune est une fille et l'enfant le plus âgé est un garçon.

Dans ces conditions, le nombre total d'enfants est : $2N$ et le nombre total de garçon dans ces familles : $2N(G_1,G_2)+N(G_1,F_2)+N(F_1,G_2)$.

Soit alors p la proportion cherchée (proportion de garçons dans ces familles de 2 enfants).

On a :

$$p = \frac{2N(G_1, G_2) + N(G_1, F_2) + N(F_1, G_2)}{2N}$$

On peut transformer ce rapport de la façon suivante :

$$\begin{aligned}
p &= \frac{2N(G_1, G_2) + N(G_1, F_2) + N(F_1, G_2)}{2N} \\
&= \frac{1}{2} \frac{N(G_1, G_2) + N(G_1, G_2) + N(G_1, F_2) + N(F_1, G_2)}{N} \\
&= \frac{1}{2} \left[\frac{N(G_1, G_2) + N(G_1, F_2)}{N} + \frac{N(G_1, G_2) + N(F_1, G_2)}{N} \right]
\end{aligned}$$

$N(G_1, G_2) + N(G_1, F_2)$ est le nombre total de familles dont l'enfant le plus jeune est un garçon. On en déduit que le rapport $\frac{N(G_1, G_2) + N(G_1, F_2)}{N}$ correspond à la proportion des familles dans lesquelles l'enfant le plus jeune est un garçon. D'après l'énoncé, cette proportion est simplement : 0,35 (35%).

De façon analogue, $\frac{N(G_1, G_2) + N(F_1, G_2)}{N}$ désigne la proportion des familles dans lesquelles l'enfant le plus âgé est un garçon. L'énoncé ne nous fournit pas cette proportion mais nous indique celle des familles dont l'enfant le plus âgé est une fille (55%). On en déduit que la proportion des familles dans lesquelles l'enfant le plus âgé est un garçon vaut $1 - 0,55$, soit 0,45.

Finalement : $p = \frac{1}{2}(0,35 + 0,45) = \frac{1}{2} 0,8 = 0,4$.

Parmi les $2N$ enfants que comptent ces familles, il y a donc 40% de garçons.

Remarque fondamentale : dans cette approche, nous n'utilisons pas le fait que dans la moitié des familles les enfants sont du même sexe.

$2^{ème}$ *approche*

Nous utilisons des notations proches de celles de la première approche. Nous noterons avec un petit *p* les probabilités calculées à partir de l'univers des familles et avec un grand *P* les probabilités calculées à partir de l'univers de l'ensemble des enfants (cet univers compte deux fois plus d'éléments que le précédent) :

- $p(G_1, G_2)$ désigne la probabilité qu'une famille comporte deux garçons. Il s'agit <u>également</u> de la probabilité $P(G_1, G_2)$ qu'un enfant choisi au hasard soit un garçon d'une famille de deux garçons : $P(G_1, G_2) = p(G_1, G_2)$;
- $p(G_1, F_2)$ désigne la probabilité qu'une famille comporte deux enfants de sexes différents, le plus jeune étant un garçon. On a ici : $P(G_1, F_2) = \frac{1}{2} p(G_1, F_2)$.
- $p(F_1, G_2)$ désigne la probabilité qu'une famille comporte deux enfants de sexes différents, le plus jeune étant une fille. On a ici : $P(F_1, G_2) = \frac{1}{2} p(F_1, G_2)$;

Nous cherchons : $P = P(G) = P(G_1, G_2) + P(G_1, F_2) + P(F_1, G_2)$ où $P(G)$ désigne la probabilité qu'un enfant choisi au hasard soit un garçon.

On a, d'après ce qui précède : $P = P(G) = p(G_1, G_2) + \frac{1}{2} p(G_1, F_2) + \frac{1}{2} p(F_1, G_2)$

D'après l'énoncé, on a :

$$p(G_1) = 0,35 = p(G_1, G_2) + p(G_1, F_2) \qquad (1)$$

$$p(F_2) = 0,55 = p(G_1, F_2) + p(F_1, F_2) \qquad (2)$$

En additionnant ces deux égalités, il vient :

$$\begin{aligned} 0,9 &= p(G_1, G_2) + p(G_1, F_2) + p(G_1, F_2) + p(F_1, F_2) \\ &= 2p(G_1, F_2) + p(G_1, G_2) + p(F_1, F_2) \\ &= 2p(G_1, F_2) + 0,5 \end{aligned}$$

D'où : $p(G_1, F_2) = 0,2$.
A partir de là, l'égalité (1) nous donne : $p(G_1, G_2) = 0,15$.

La somme $p(G_1, F_2) + p(F_1, G_2)$ est égale à la probabilité que les enfants d'une famille donnée ne soient pas du même sexe. D'après l'énoncé, on a :

$$p(G_1, F_2) + p(F_1, G_2) = 1 - 0,5 = 0,5$$

Finalement :

$$\begin{aligned} P &= p\left(G_1, G_2\right) + \frac{1}{2} p\left(G_1, F_2\right) + \frac{1}{2} p\left(F_1, G_2\right) \\ &= 0,15 + \frac{1}{2} 0,5 \\ &= 0,4 \end{aligned}$$

On retrouve le résultat obtenu précédemment.

Résultat final

Parmi les enfants des familles considérées, il y a 40% de garçons.