**Objectif :** comprendre les notions de base d'une distribution discrète, savoir calculer les espérances et variances d'une distribution discrète.

**Ex 1**. L'entreprise Time Warner Cable fournit des services de télévision et d'internet. Sur la base des adhésions passées, elle estime la probabilité du nombre d'adhésions auxquelles elle peut s'attendre pour l'année prochaine comme suit :

| Nombre d'adhésions futures | Probabilité estimée |
|---|---|
| 100 000 | 0,10 |
| 200 000 | 0,20 |
| 300 000 | 0,25 |
| 400 000 | 0,30 |
| 500 000 | 0,10 |
| 600 000 | 0,05 |

1) Quelle est la probabilité que l'entreprise obtienne plus de 400 000 nouveaux abonnés ?

Puisque l'on demande la probabilité que Time Warner Cable obtienne plus de 400 000 abonnés (signifiant strictement plus de 400 000), on ajoute donc les probabilités de 500 000 et 600 000, soit : $\mathbb{P}(500\,000) + \mathbb{P}(600\,000) = 0{,}10 + 0{,}05 = 0{,}15$. Ainsi, la probabilité que l'entreprise obtienne plus de 400 000 nouveaux abonnés est de 15%.

2) Quelle est la probabilité que l'entreprise obtienne moins de 200 000 nouveaux abonnés ?

Même chose ici, on demande la probabilité que Time Warner Cable obtienne moins de 200 000 abonnés (signifiant strictement moins de 200 000), on ne prend donc que la probabilité de 100 000, soit : $\mathbb{P}(100\,000) = 0{,}10$. Ainsi, la probabilité que l'entreprise obtienne moins de 200 000 nouveaux abonnés est de 10%.

**Ex 2**. Une enquête menée auprès de chômeurs dans une région américaine à la fin décembre 2009 a permis de recueillir les données suivantes sur le nombre de mois passés au chômage par ces derniers :

| Mois de chômage | Nombre de chômeurs |
|---|---|
| 1 | 1029 |
| 2 | 1686 |
| 3 | 2269 |
| 4 | 2675 |
| 5 | 3487 |
| 6 | 4652 |
| 7 | 4145 |
| 8 | 3587 |
| 9 | 2325 |
| 10 | 1120 |

1) On cherche à connaître les probabilités associées aux différentes durées de chômage. Comment doit-on procéder ?

Pour obtenir ces probabilités (c'est-à-dire donner la distribution de la variable moins de chômage, ou encore sa loi), il suffit de diviser le nombre d'occurrences de chaque modalité de la variable par le total des occurrences, pour toutes les durées de chômage. Ici, le total des chômeurs représente 26 975. On obtient ainsi :

| Mois de chômage | Nombre de chômeurs | Probabilité |
|---|---|---|
| 1 | 1029 | $1029/26975 = 0{,}04$ |
| 2 | 1686 | $1686/26975 = 0{,}06$ |
| 3 | 2269 | $2269/26975 = 0{,}08$ |
| 4 | 2675 | $2675/26975 = 0{,}10$ |
| 5 | 3487 | $3487/26975 = 0{,}13$ |
| 6 | 4652 | $4652/26975 = 0{,}17$ |
| 7 | 4145 | $4145/26975 = 0{,}15$ |
| 8 | 3587 | $3587/26975 = 0{,}13$ |
| 9 | 2325 | $2325/26975 = 0{,}09$ |
| 10 | 1120 | $1120/26975 = 0{,}05$ |

2) En déduire la probabilité qu'une personne sélectionnée aléatoirement soit au chômage
   a. Pendant 2 mois au plus

On demande la probabilité qu'une personne sélectionnée aléatoirement soit au chômage 2 mois au plus (signifiant 2 mois ou moins). On ajoute donc les probabilités de 1 et 2 mois, soit : $\mathbb{P}(1) + \mathbb{P}(2) = 0{,}04 + 0{,}06 = 0{,}10$. Ainsi, la probabilité qu'une personne sélectionnée aléatoirement soit au chômage 2 mois au plus est de 10%.

   b. Pendant plus de 2 mois

De la même manière, plus de 2 moins, soit entre 3 et 10 mois, correspond donc à :

$$\sum_{3}^{10} \mathbb{P}(x_i) = 1 - \big(\mathbb{P}(1) + \mathbb{P}(2)\big) = 0{,}90$$

La probabilité qu'une personne sélectionnée aléatoirement soit au chômage plus de 2 mois est de 90%.

   c. Pendant plus de 6 mois

Plus de 6 mois correspond à entre 7 et 10 mois, ainsi :

$$\sum_{7}^{10} \mathbb{P}(x_i) = 0{,}15 + 0{,}13 + 0{,}09 + 0{,}05 = 0{,}42$$

La probabilité qu'une personne sélectionnée aléatoirement soit au chômage plus de 6 mois est de 42%.

**Ex 3**. La demande pour un produit de l'entreprise Carolina Industries varie considérablement d'un mois à l'autre. Basée sur les données des deux dernières années, l'entreprise a estimé les probabilités de demande mensuelle suivantes :

| Demande mensuelle (en unités de produit) | Probabilité |
|---|---|
| 300 | 0,20 |
| 400 | 0,30 |
| 500 | 0,35 |
| 600 | 0,15 |

1) Si cette entreprise base sa production sur l'espérance mathématique de la demande mensuelle, quelle quantité doit-elle produire par mois ?

Pour le savoir, on doit donc calculer l'espérance mathématique de la demande. Ceci revient à faire la somme des produits des demandes par leur probabilité respective. Soient $x_i$ la variable aléatoire représentant le niveau de demande mensuel. On a :

$$\sum_{i=1}^{4} x_i \times \mathbb{P}(x_i) = (300 \times 0{,}20) + (400 \times 0{,}30) + (500 \times 0{,}35) + (600 \times 0{,}15) = 445$$

Si cette entreprise base sa production sur l'espérance mathématique de la demande mensuelle, elle devra donc produire 445 unités par mois.

2) Supposons que chaque unité est vendue à 70$ pour un coût de production unitaire de 50$. À combien se montera le profit de l'entreprise sur le mois si son volume de production ce mois donné est celui établi en 1) et que la demande effective pour le produit vaut 300 ?

On remarque que si l'entreprise se base sur le critère d'espérance mathématique pour établir son niveau mensuel de production, elle tendra à le surestimer pour le mois en question. Ici, sur 445 unités produites, elle n'en écoulera que 300, soit 145 de moins. Son profit Π du mois se montera donc à :

$$\Pi = \big(300 \times (70 - 50)\big) - (145 \times 50) = 6000 - 7250 = -1250$$

**Ex 4**. L'entreprise J. R. Computer envisage d'agrandir son usine pour permettre la production d'un nouvel ordinateur mais hésite sur l'ampleur de l'agrandissement car la demande pour ce nouveau produit est incertaine. Le bureau d'études estime les probabilités pour une demande faible, moyenne ou élevée à respectivement 0,20, 0,50 et 0,30. Il établit également des prévisions sur les bénéfices associés selon les deux projets en lice. Soit $X$ le bénéfice annuel en milliers de dollars attendu pour le projet le moins ambitieux et $Y$ celui attendu pour le projet le plus ambitieux.

| | $X$ | $Y$ |
|---|---|---|
| Demande faible | 50 | 0 |
| Demande moyenne | 150 | 100 |
| Demande élevée | 200 | 300 |

1) Calculer l'espérance mathématique du bénéfice associé à chacun des projets. Quel est le projet qui maximise l'espérance de bénéfice ?

On calcule donc les espérances mathématiques de gain pour les projets $X$ et $Y$ :

$$\mathbb{E}(X) = \sum_{i=1}^{3} X_i \mathbb{P}(X_i) = (50 \times 0{,}20) + (150 \times 0{,}50) + (200 \times 0{,}30) = 145$$

$$\mathbb{E}(Y) = \sum_{i=1}^{3} Y_i \mathbb{P}(Y_i) = (0 \times 0{,}20) + (100 \times 0{,}50) + (300 \times 0{,}30) = 140$$

On constate qu'en termes d'espérance, le projet $X$ est *a priori* le plus lucratif.

2) Calculer la variance du bénéfice pour les deux projets. Quelle décision faut-il prendre si l'objectif est de minimiser l'incertitude (le risque) ?

Calculons la variance du bénéfice des deux projets (on se sert de la formule de König-Huygens pour simplifier les calculs :

$$\mathbb{V}(X) = \mathbb{E}\left(\left(X - \mathbb{E}(X)\right)^2\right) = \mathbb{E}(X^2) - \left(\mathbb{E}(X)\right)^2$$

$$\mathbb{E}(X^2) = \sum_{i=1}^{3} X_i^2 \mathbb{P}(X_i) = (50^2 \times 0{,}20) + (150^2 \times 0{,}50) + (200^2 \times 0{,}30) = 23750$$

$$\left(\mathbb{E}(X)\right)^2 = 145^2 = 21025$$

Ainsi, $\mathbb{V}(X) = 23750 - 21025 = 2725$.

$$\mathbb{V}(Y) = \mathbb{E}\left(\left(Y - \mathbb{E}(Y)\right)^2\right) = \mathbb{E}(Y^2) - \left(\mathbb{E}(Y)\right)^2$$

$$\mathbb{E}(Y^2) = \sum_{i=1}^{3} Y_i^2 \mathbb{P}(Y_i) = (0^2 \times 0{,}20) + (100^2 \times 0{,}50) + (300^2 \times 0{,}30) = 32000$$

$$\left(\mathbb{E}(Y)\right)^2 = 140^2 = 19600$$

Ainsi, $\mathbb{V}(Y) = 32000 - 19600 = 12400$.

On constate facilement que le projet $Y$ est de loin le plus risqué, la variabilité de ses gains étant très élevée. Si l'on souhaite minimiser les risques, on se portera sur le projet $X$. De plus, les gains moyens sont inférieurs pour le projet $Y$ de toutes façons…

**Ex 5**. On suppose que le nombre de livres achetés par un client dans une librairie est une variable aléatoire notée $X$. On connaît :

$$\mathbb{P}(0 \leq X \leq 1) = 0{,}6 \,;\, \mathbb{P}(1 \leq X \leq 2) = 0{,}5 \,;\, \mathbb{P}(0 \leq X < 3) = 0{,}8 \,;$$

$$\mathbb{P}(X = 3) = \mathbb{P}(X \geq 4) \,;\, \mathbb{P}(X = 4) = \mathbb{P}(X \geq 5)$$

Calculer $\mathbb{P}(X = i)$ pour $i \in \{0, 1, 2, 3, 4\}$ uniquement.

Des relations présentées par l'énoncé, on peut déduire :

$$(1) :\; \mathbb{P}(0 \leq X \leq 1) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) = 0{,}6$$

$$(2) :\; \mathbb{P}(1 \leq X \leq 2) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = 0{,}5$$

$$(3) :\; \mathbb{P}(0 \leq X < 3) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = 0{,}8$$

Donc, par soustraction, on a facilement :

$$(3) - (1) :\; \mathbb{P}(X = 2) = 0{,}8 - 0{,}6 = 0{,}2$$

$$(3) - (2) :\; \mathbb{P}(X = 0) = 0{,}8 - 0{,}5 = 0{,}3$$

De plus :

$$\mathbb{P}(0 \leq X < 3) - \mathbb{P}(X = 0) - \mathbb{P}(X = 2) = \mathbb{P}(X = 1) = 0{,}8 - 0{,}3 - 0{,}2 = 0{,}3$$

On en déduit alors que, puisque $\mathbb{P}(0 \leq X < 3) = 0{,}8$ et que $\mathbb{P}(X = 3) = \mathbb{P}(X \geq 4)$, on a :

$$\mathbb{P}(X \geq 3) = 1 - \mathbb{P}(0 \leq X < 3) = 1 - 0{,}8 = 0{,}2 \text{ et } \mathbb{P}(X = 3) = \mathbb{P}(X \geq 4) = \frac{0{,}2}{2} = 0{,}1$$

De même, puisque $\mathbb{P}(X \geq 4) = 0{,}1$ et que $\mathbb{P}(X = 4) = \mathbb{P}(X \geq 5)$, alors :

$$\mathbb{P}(X = 4) = \mathbb{P}(X \geq 5) = \frac{0{,}1}{2} = 0{,}05$$

**Ex 6**. On lance 1 fois un dé équilibré. Soit $X$ la variable aléatoire égale au nombre de points qui figurent sur la face du dé.

On a un dé équilibré ici, donc l'univers consiste en l'ensemble des 6 faces du dé, soit :

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

De plus, puisque le dé est équilibré, tous les éléments de $\Omega$ sont équiprobables. On a donc :

$$\forall \omega \in \Omega, \mathbb{P}(\omega) = \frac{1}{|\Omega|} = \frac{1}{6}$$

1) Quelle est la loi de $X$ ? Quelle est la valeur moyenne de $X$ ?

La loi d'une variable aléatoire discrète correspond à la probabilité d'occurrence de chacune de ses modalités. Ici, on a très facilement :

$$\mathbb{P}(X_i) = \frac{1}{6}, \forall i \in \Omega$$

On peut donc écrire la loi le $X$ telle que :

$$\mathbb{P}_X(i) = \frac{1}{6}, \forall i \in \Omega$$

Pour la valeur moyenne de $X$, on passe par son espérance :

$$\mathbb{E}(X) = \sum_{i=1}^{6} X_i \mathbb{P}(X_i) = \left(1 \times \frac{1}{6}\right) + \left(2 \times \frac{1}{6}\right) + \left(3 \times \frac{1}{6}\right) + \left(4 \times \frac{1}{6}\right) + \left(5 \times \frac{1}{6}\right) + \left(6 \times \frac{1}{6}\right) = \frac{21}{6}$$

$$= 3{,}5$$

2) On suppose qu'on reçoit 15 euros si on obtient « 1 », 5 euros si on obtient « 5 » ou « 6 » et rien si on obtient « 2 », « 3 » ou « 4 ». Si on nomme $Y$ la variable aléatoire égale au gain de ce jeu, quelle est la loi de $Y$ ? Que vaut le gain moyen ?

La loi de $Y$ correspond aux probabilités associées aux valeurs qu'elle peut prendre. Ici, l'ensemble des valeurs de $Y$ (aussi appelé le support de $Y$), est simplement : $Y(\Omega) = \{15, 5, 0\}$.

$$\mathbb{P}_Y(15) = \mathbb{P}(Y = 15) = \mathbb{P}(X = 1) = \frac{1}{6}$$

$$\mathbb{P}_Y(5) = \mathbb{P}(Y = 5) = \mathbb{P}(X = 1) + \mathbb{P}(X = 6) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

$$\mathbb{P}_Y(0) = \mathbb{P}(Y = 0) = \mathbb{P}(X = 2) + \mathbb{P}(X = 3) + \mathbb{P}(X = 4) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$$

Ainsi, le gain moyen à ce jeu est donné par :

$$\mathbb{E}(Y) = \left(\frac{1}{6} \times 15\right) + \left(5 \times \frac{1}{3}\right) + \left(0 \times \frac{1}{2}\right) = \frac{25}{6} = 4{,}16$$

Le gain moyen est donc de 4,16€.

3) On suppose maintenant qu'on reçoit 25 euros pour « 1 » et rien sinon. Préférez-vous jouer au jeu précédent ou à celui-ci ? Pourquoi ?

Soit $Z$ la variable aléatoire associée aux gains de ce deuxième jeu. La loi de $Z$ correspond aux probabilités associées aux valeurs qu'elle peut prendre. Ici, l'ensemble des valeurs de $Z$ (aussi appelé le support de $Z$), est simplement : $Z(\Omega) = \{25, 0\}$.

$$\mathbb{P}_Z(25) = \mathbb{P}(Z = 25) = \mathbb{P}(X = 1) = \frac{1}{6}$$

$$\mathbb{P}_Z(0) = 1 - \mathbb{P}(Z = 25) = 1 - \mathbb{P}(X = 1) = \frac{5}{6}$$

Ainsi, le gain moyen à ce jeu est donné par :

$$\mathbb{E}(Z) = \left(\frac{1}{6} \times 25\right) + \left(0 \times \frac{5}{6}\right) = \frac{25}{6} = 4{,}16$$

Le gain moyen est donc aussi de 4,16€. Les deux jeux sont donc équivalents en moyenne. Néanmoins, si l'on est averse au risque, on préfèrera sans doute le premier jeu, dans la mesure où sa variance est moins élevée, et qu'un gain quelconque est plus probable.