Variance pour les lois discrètes

Licence Sciences pour la Santé

Intervenant : Mathieu Fauvernier

Espérance d'une variable aléatoire discrète

- L'espérance d'une v.a.r correspond à sa valeur moyenne

$$E[X] = \sum_{i=1}^{+\infty} x_i P(X = x_i)$$

Espérance d’une variable aléatoire discrète

- Soit une fonction h, comment définir $E[h(X)]$?

Espérance d'une variable aléatoire discrète

- Exemple : soit X la variable aléatoire telle que les valeurs possibles et probabilités associées sont données ci-dessous

x	-2,5	0	1	2
P(X=x)	0,2	0,4	0,3	0,1

- Calculer E(X) et $E(X^2)$

Espérance d’une variable aléatoire discrète

- Exemple : soit X la variable aléatoire telle que les valeurs possibles et probabilités associées sont données ci-dessous

x	**-2,5**	**0**	**1**	**2**
P(X=x)	0,2	0,4	0,3	0,1

- $E(X) = -2,5 * 0,2 + 0 * 0,4 + 1 * 0,3 + 2 * 0,1 = 0$

Espérance d’une variable aléatoire discrète

- Exemple : soit X la variable aléatoire telle que les valeurs possibles et probabilités associées sont données ci-dessous

x	**-2,5**	**0**	**1**	**2**
x^2	6,25	0	1	4
P(X=x)	0,2	0,4	0,3	0,1

- $E(X^2) = ?$

Espérance d'une variable aléatoire discrète

- Exemple : soit X la variable aléatoire telle que les valeurs possibles et probabilités associées sont données ci-dessous

x	-2,5	0	1	2
x^2	6,25	0	1	4
P(X=x)	0,2	0,4	0,3	0,1

- $E(X^2) = 6{,}25 * 0{,}2 + 0 * 0{,}4 + 1 * 0{,}3 + 4 * 0{,}1 = 1{,}95$

Espérance d'une variable aléatoire discrète

- Soit une fonction h, comment définir $E[h(X)]$?

$$E[\mathrm{h}(X)] = \sum_{i=1}^{+\infty} h(x_i)P(X = x_i)$$

Variance d'une variable aléatoire

La variance d'une v.a.r mesure la dispersion autour de la moyenne

$$\mathrm{Var}[X] = \mathrm{E}[(X - E[X])^2] = E[X^2] - E[X]^2$$

L'écart-type correspond à la racine carrée de la variance

$$\sigma[X] = \sqrt{Var[X]}$$

Variance d'une variable aléatoire

La variance d'une variable aléatoire mesure la dispersion autour de la moyenne

$$\mathrm{Var}[X] = \mathrm{E}[(X - E[X])^2] = E[X^2] - E[X]^2$$

$$E[X^2] = \sum_{i=1}^{+\infty} {x_i}^2 P(X = x_i)$$

$$\mathrm{Var}[X] = \sum_{i=1}^{+\infty} (x_i - E(X))^2 P(X = x_i)$$

Problème 1

- Naissances multiples

x	1	2	3	4	5+
P(X=x)	96%	3%	0,9%	0,07%	0,03%

Le ministère des Solidarités et de la Santé estime le coût annuel moyen d'un enfant à 9 000 €.

Quelle est la loi du coût moyen, son espérance et sa variance ?

Problème 1

- Naissances multiples

x	1	2	3	4	5+
coût	9k	18k	27k	36k	45k
P(X=x)	96%	3%	0,9%	0,07%	0,03%

$E(Coût) = E(9000 * X) = 9000 * E(X) = 9461{,}7$ avec

$$E(X) = 1 * 0{,}96 + 2 * 0{,}03 + 3 * 0{,}009 + 4 * 0{,}0007 + 5 * 0{,}0003 = 1{,}0513$$

Problème 1

- Naissances multiples

x	1	2	3	4	5+
coût	9k	18k	27k	36k	45k
P(X=x)	96%	3%	0,9%	0,07%	0,03%

$Var(Cout) = Var(9000 * X) = 9000^2 * Var(X) = 6031933$ avec

$$Var(X) = (1 - 1{,}05)^2 * 0{,}96 + (2 - 1{,}05)^2 * 0{,}03 + (3 - 1{,}05)^2 * 0{,}009 + (4 - 1{,}05)^2 * 0{,}0007 + (5 - 1{,}05)^2 * 0{,}0003 \approx 0{,}0745$$

Problème 1

- Naissances multiples

x	1	2	3	4	5+
coût	9k	18k	27k	36k	45k
P(X=x)	96%	3%	0,9%	0,07%	0,03%

$Var(Cout) = Var(9000 * X) = 9000^2 * Var(X) = 6031933$ avec

L'unité de la variance est en euros^2 donc pas super interprétable !

Ecart-type = $\sqrt{6031933} = 2456$ euros

Problème 2

- On s'intéresse à la pression artérielle systolique prédite à partir de l'âge. On dispose d'un groupe de patientes dont la distribution d'âge est la suivante (on donne les âges médians de 5 classes d'âges)

age	**20**	**30**	**40**	**50**	**60**
P(Age=age)	5%	10%	20%	30%	35%

- Pour un âge donné a, la pression est donnée par (en mmHg) :

$$Pre(a) = 0{,}01 * a^2 + 0{,}05 * a + 107$$

Calculer E(Age), Var(Age), ainsi que E(Pre) et Var(Pre)

Problème 2

- On s'intéresse à la pression artérielle systolique prédite à partir de l'âge. On dispose d'un groupe de patientes dont la distribution d'âge est la suivante (on donne les âges médians de 5 classes d'âges)

age	20	30	40	50	60
P(Age=age)	5%	10%	20%	30%	35%

$$E(Age) = 20 * 0{,}05 + 30 * 0{,}1 + 40 * 0{,}2 + 50 * 0{,}3 + 60 * 0{,}35 = 48$$

$$\begin{aligned}&\mathrm{Var}(Age)\\ &= (20-48)^2 * 0{,}05 + (30-48)^2 * 0{,}1 + (40-48)^2 * 0{,}2\\ &+ (50-48)^2 * 0{,}3 + (60-48)^2 * 0{,}35 = 136\end{aligned}$$

Problème 2

- On s'intéresse à la pression artérielle systolique prédite à partir de l'âge. On dispose d'un groupe de patientes dont la distribution d'âge est la suivante (on donne les âges médians de 5 classes d'âges)

age	20	30	40	50	60
P(Age=age)	5%	10%	20%	30%	35%

Sachant que $\mathrm{E}(Age) = 48$, et que

$$Pre(a) = 0{,}01 * a^2 + 0{,}05 * a + 107$$

A-t-on $\mathrm{E}(Pre) = Pre(48) = 132{,}44$?

Problème 2

- On s'intéresse à la pression artérielle systolique prédite à partir de l'âge. On dispose d'un groupe de patientes dont la distribution d'âge est la suivante (on donne les âges médians de 5 classes d'âges)

age	20	30	40	50	60
Pre	112	117,5	125	134,5	146
P(Age=age)	5%	10%	20%	30%	35%

$$E(Pre)$$
$$= 112 * 0{,}05 + 117{,}5 * 0{,}1 + 125 * 0{,}2 + 134{,}5 * 0{,}3 + 146 * 0{,}35$$
$$= 133{,}8 \neq Pre(48)$$

Problème 2

- On s'intéresse à la pression artérielle systolique prédite à partir de l'âge. On dispose d'un groupe de patientes dont la distribution d'âge est la suivante (on donne les âges médians de 5 classes d'âges)

age	20	30	40	50	60
Pre	112	117,5	125	134,5	146
P(Age=age)	5%	10%	20%	30%	35%

$$Var(Pre) = (112 - 133{,}8)^2 * 0{,}05 + (117{,}5 - 133{,}8)^2 * 0{,}1 + (125 - 133{,}8)^2 * 0{,}2 + (134{,}5 - 133{,}8)^2 * 0{,}3 + (146 - 133{,}8)^2 * 0{,}35 = 118{,}06$$

Problème 2

- On s'intéresse à la pression artérielle systolique prédite à partir de l'âge. On dispose d'un groupe de patientes dont la distribution d'âge est la suivante (on donne les âges médians de 5 classes d'âges)

age	20	30	40	50	60
Pre	112	117,5	125	134,5	146
P(Age=age)	5%	10%	20%	30%	35%

$$Var(Pre) = 118{,}06$$

$$Ecart\_type(Pre) = 10{,}87$$

Problème 3

Des patients subissent une opération chirurgicale nécessitant l'administration préalable d'un réactif. La distribution des durées d'opérations est donnée ci-dessous

durée	1	2	3	4	5
P(Durée=durée)	10%	40%	30%	15%	5%

Sachant que la quantité de réactif présente dans l'organisme en fonction du temps est

$$R(t) = -t^2 + 6 * t$$

Calculer $E(R(Durée))$ et $\mathrm{Var}(R(Durée))$

Problème 3

Des patients subissent une opération chirurgicale nécessitant l'administration préalable d'un réactif. La distribution des durées d'opérations est donnée ci-dessous

durée	1	2	3	4	5
R(durée)	5	8	9	8	5
P(Durée=durée)	10%	40%	30%	15%	5%

$$E\big(R(Durée)\big) = 5 * 0{,}1 + 8 * 0{,}4 + 9 * 0{,}3 + 8 * 0{,}15 + 5 * 0{,}05 = 7{,}85$$

Problème 3

Des patients subissent une opération chirurgicale nécessitant l'administration préalable d'un réactif. La distribution des durées d'opérations est donnée ci-dessous

durée	1	2	3	4	5
R(durée)	5	8	9	8	5
P(Durée=durée)	10%	40%	30%	15%	5%

$$Var\big(R(Durée)\big) = (5-7{,}85)^2 * 0{,}1 + (8-7{,}85)^2 * 0{,}4 + (9-7{,}85)^2 * 0{,}3 + (8-7{,}85)^2 * 0{,}15 + (5-7{,}85)^2 * 0{,}05 = 1{,}6275$$

Problème 3

Des patients subissent une opération chirurgicale nécessitant l'administration préalable d'un réactif. La distribution des durées d'opérations est donnée ci-dessous

durée	1	2	3	4	5
R(durée)	5	8	9	8	5
P(Durée=durée)	10%	40%	30%	15%	5%

Sachant que l'on doit garder les patients en salles de réveil tant qu'il reste du réactif dans leur organisme. Combien de temps, en moyenne, un patient attendra-t-il en salle de réveil ?

Problème 3

durée	1	2	3	4	5
R(durée)	5	8	9	8	5
P(Durée=durée)	10%	40%	30%	15%	5%

La quantité de réactif présente dans l'organisme en fonction du temps est

$$R(t) = -t^2 + 6 * t$$

Cette quantité est nulle lorsque $R(t) = 0 \Leftrightarrow t = 6$ ou $t = 0$

Problème 3

Durée	1	2	3	4	5
Temps d'attente	5	4	3	2	1
P(Durée=durée)	10%	40%	30%	15%	5%

$$E(Temps\ attente) = 5 * 0{,}1 + 4 * 0{,}4 + 3 * 0{,}3 + 2 * 0{,}15 + 1 * 0{,}05$$
$$= 0{,}5 + 1{,}6 + 0{,}9 + 0{,}3 + 0{,}05 = 3{,}35$$