GM – Probabilités et Statistique

http://moodle.epfl.ch/course/view.php?id=18431

Cours 4

- Révision : variables aléatoires
- Espérance et variance
- Révision : loi binomiale
- Loi de Poisson
- Approximation de la loi binomiale par la loi de Poisson

Révision : VAs discrètes

- **Variable Aléatoire (VA)** : une fonction réelle définie sur l'ensemble fondamental
  - VA : MAJUSCULES ; valeur spécifique : miniscules
- **VA discrète** :
  1. loi de probabilité : $p(x) = P(X = x)$
  2. fonction de répartition : $F(x) = P(X \leq x) = \sum_{i \leq x} p(i)$
- **Resolution des problèmes avec VAs**
  1. *Identifier la VA*
  2. Déterminer la *distribution* (loi) de la VA
  3. *Traduire* la question
  4. *Répondre* à la question

Espérance

- Pour une VA discrète X de loi $p(x)$, on définit l'**espérance** (ou la **moyenne**) par :

$$E[X] = \sum_{\substack{\text{toutes} \\ \text{valeurs } x}} xp(x)$$

- Donc c'est la *moyenne pondérée* des valeurs possibles de X, où les poids sont $P(X = x)$
- C'est également possible à calculer l'espérance d'une *fonction* de la VA (discrète) X (disons $g(X)$) dans la même manière
- $g(X)$ elle aussi est une VA discrète, donc pour calculer $E[g(X)]$ il suffira de trouver sa loi (distribution) $p(g(x))$
- On devrait pouvoir déduire la distribution de celle de X

Exemple

Exemple 4.1 Soit $X =$ la somme des nombres de deux dès lancés indépendamment.

$E[X] =$

x	2	3	4	5	6	7	8	9	10	11	12	
$p(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	
$F(x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$	$[= 1]$

[Grâce à la symetrie de la loi de X, ceci est ce qu'on aurait deviné, sans bénéfice d'un cours de probabilités ! !]

Exemple

$\boxed{\textbf{Exemple 4.2}}$ Calculer $E[X^2]$ pour la loi de X suivante :

$$P(X=-1)=0.2 \qquad P(X=0)=0.5 \qquad P(X=1)=0.3.$$

Solution : On définit une nouvelle VA $Y=X^2$. Trouvons la distribution de Y directement :

$$P(Y=1)=P(X=-1)+P(X=1)=0.2+0.3=0.5$$
$$P(Y=0)=P(X=0)=0.5.$$

Donc $E[X^2]=E[Y]=1(0.5)+0(0.5)=\underline{\underline{0.5}}$

A noter : $(E[X])^2 \boxed{\neq} E[X^2]$

$E[g(X)]$ encore une fois

- *Autre façon* de voir $E[g(X)]$:
  en notant que $g(X) = g(x)$ lorsque $X = x$, il est raisonnable de penser que $E[g(X)]$ puisse être *la moyenne pondérée* des valeurs $g(x)$, poids $P(X = x)$
- **Théorème :** Si X est une VA discrète pouvant prendre ses valeurs parmi les valeurs x_i, $i \geq 1$, avec des probabilités respectives $p(x_i)$, alors pour toute fonction réelle g on a

$$E[g(X)] = \sum_i g(x_i)p(x_i)$$

- Pour toute paire (a, b) de constantes, $E[aX + b] = aE[X] + b$

Résumés d'une distribution

- Une VA X et sa fonction de répartition F (ou loi $p(x)$) étant données, il serait utile de *résumer les propriétés* de F en deux ou trois mesures
- Une telle mesure est donné par $E[X]$, l'espérance de X, qui nous dit quelque chose sur la valeur 'centrale' de la distribution
- Cependent, elle ne nous dit rien des *variations* de X autour de l'espérance

Exemple

- Considerons les VAs W, Y, et Z :

$$
\begin{aligned}
W &= 0 \\
Y &= \begin{cases} -1 \text{ avec probabilité } \frac{1}{2} \\ +1 \text{ avec probabilité } \frac{1}{2} \end{cases} \\
Z &= \begin{cases} -100 \text{ avec probabilité } \frac{1}{2} \\ +100 \text{ avec probabilité } \frac{1}{2} \end{cases}
\end{aligned}
$$

- Toutes les 3 ont la même espérance ($=$ *? ?*), *mais les écarts entre les différentes valeurs* de Y sont plus grandes que celles de W, et plus petites que celles de Z

Variance et écart-type

- Comme on s'attend à voir toute variable X prendre ses valeurs autour de son espérance $E[X]$, il paraît raisonnable de mesurer les variations en considérant l'écart moyen entre X et $E[X]$, $E[\mid X - \mu \mid]$, où $\mu = E[X]$
- Il est plus facile techniquement (en maths) de considerer le moyen du *carré* de l'écart entre X et sa espérance $E[X]$
- Pour la VA X avec espérance μ, on définit la **variance** de X :

$$Var(X) = E[(X - \mu)^2]$$

- On peut établir une *formule alternative* pour le calcul de $Var(X)$ (plus commode dans la pratique) :

$$Var(X) = E[X^2] - (E[X])^2$$

- L'**écart-type** de X (σ) est la racine carré de $Var(X)$:

$$\sigma = \sqrt{Var(X)}$$

Variance d'une fonction linéaire d'une VA X

- Pour toute paire (a, b) de constantes,

$$Var(aX + b) = a^2 Var(X)$$

- Facile à démontrer :

$$\begin{aligned} Var(aX + b) &= E[(aX + b) - E(aX + b)]^2 \\ &= E[aX + b - (aE[X] + b)]^2 \\ &= E[aX - aE[X]]^2 \\ &= E[a^2(X - E[X])^2]] \\ &= a^2 E[(X - \mu)^2] \\ &= a^2 Var(X) \end{aligned}$$

- Donc pour une fonction linéaire de X, on a :

$$SD(aX + b) = \mid a \mid SD(X)$$

- ('SD' = 'écart-type' = 'standard deviation' en anglais)

VA de Bernoulli

- Une VA de *Bernoulli* prend les valeurs 0 et 1
- Sa loi de probabilité est :

x	0	1
$p(x)$	$(1-p)$	p

- Utilisée dans la modelisation des problèmes ayant 2 résultats possibles : pile/face ; oui/non ; succès/échec ; *etc.*
- Pour une VA de Bernoulli X :

$$\begin{aligned} E(X) &= 0 \times (1-p) + 1 \times p = \mathbf{p}; \\ Var(X) &= E[X^2] - (E[X])^2 = [0^2 \times (1-p) + 1^2 \times p] - p^2 \\ &= p - p^2 = \mathbf{p(1\text{-}p)} \end{aligned}$$

VA binomiale

- Il y a *4 conditions* à satisfaire :
  - nombre *fixe* (pas aléatoire) d'épreuves n
  - 2 résultats possibles de chaque épreuve : *soit 1 soit 0*
  - *la même probabilité* p pour chaque épreuve que le résultat soit 1
  - les épreuves sont *indépendantes*
- Donc si $X \sim Bin(n, p)$, la loi de probabilité est :

$$\boxed{P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \; ; \quad \underline{\underline{x = 0, 1, \ldots, n}}}$$

Espérance et variance d'une VA binomiale

- Je lance une pièce équitable 10 fois (indépendamment). Combien de piles attendez-vous ? ?
- Je lance un dé équitable 60 fois (indépendamment). Combien de '1' attendez-vous ? ?
- L'*espérance* de $X \sim Bin(n, p)$ est $E[X] =$ **np** (en accord avec l'intuition)
- La *variance* est $Var(X) =$ **np(1-p)** (ce qui est moins intuitif ! !)

Démonstration : $E[X]$ **ne sera pas examiné**

$$E[X] = \sum_{k=0}^{n} kP(X = k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=\mathbf{1}}^{n} k \cdot \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \cdot \frac{n \cdot (n-1)!}{k \cdot (k-1)!(n-k)!} p \cdot p^{k-1} (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k}$$

Renommons les indices de la somme : $m = n - 1; i = k - 1$. Cela ne change pas la somme, mais c'est plus facile à voir qu'*on fait la sommation pour toutes valeurs possible d'une VA binomiale :*

$$= np \sum_{i=0}^{m} \frac{m!}{i!(m-i)!} p^i (1-p)^{m-i} = np \cdot 1 = \underline{\boldsymbol{np}}$$

(La dérivation de la variance est similaire.)

VA discréte : un autre exemple

Exemple 4.3 La loi de X est : $p(i) = c\lambda^i/i!$, $i = 0, 1, 2, \ldots$, où λ est un réel positif. Trouver :

(a) $P(X = 0)$ **(b)** $P(X > 2)$

Solution

- Au début, on cherche la valeur c ; puisque $\sum_{i=0}^{\infty} p(i) = 1$:

$$c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = 1$$

- En rappelant (! !) que $\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x$, on a $ce^{\lambda} = 1 \Rightarrow \boxed{c = e^{-\lambda}}$
- Cette VA est une *VA de Poisson*

Solution, cont.

Donc :

(a) $P(X=0) = e^{-\lambda}\lambda^0/0! = \boxed{e^{-\lambda}}$ [substitution ($i = 0$) en $p(i)$]

(b)

$$\begin{aligned} P(X>2) &= 1 - P(X \leq 2) && \text{[prob. év. complémentaire]} \\ &= 1 - P(X=0) - P(X=1) - P(X=2) && \text{[évs. ME]} \\ &= \boxed{1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2 e^{-\lambda}}{2}} && \text{[substitution en } p(i)\text{]} \end{aligned}$$

PAUSE

Approximation de la loi bionomiale par une loi de Poisson

- On peut utiliser un VA de Poisson *pour approximer une VA binomiale* de paramètres (n, p) pour autant que n soit grand et p assez petit pour que np soit d'ordre de grandeur moyen
- Soit $X \sim Bin(n, p)$ et $\lambda = np$; alors

$$\begin{aligned} P(X = i) &= \frac{n!}{(n-i)!i!}\, p^i (1-p)^{n-i} = \frac{n!}{(n-i)!i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\ &= \frac{n(n-1)\cdots(n-i+1)}{n^i} \cdot \frac{\lambda^i}{i!} \cdot \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i} \end{aligned}$$

- Pour n grand et λ modéré,

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda} \qquad \frac{n(n-1)\cdots(n-i+1)}{n^i} \approx 1 \qquad \left(1 - \frac{\lambda}{n}\right)^i \approx 1$$

- $\Rightarrow$ Donc $\boxed{P(X = i) \approx e^{-\lambda}\lambda^i/i! \;;\quad \underline{\underline{x = 0, 1, 2, \ldots}}}$

Applications de la loi de Poisson

- Quelques exemples de VA $\sim$ Poisson :
  - le nombre de coquilles par page d'un livre
  - le nombre de faux numéros téléphoniques composés en un jour
  - le nombre de clients pénétrant dans un bureau de poste donné en l'espace d'un jour
  - le nombre de particules α émises par une substance radioactive pendant un certain laps de temps
  - le nombre de colonies bactériennes qui se multiplient dans une boîte de Pétri en milieu nutritif favorable
  - le nombre de mutants résultant d'une expérience
- Dans chacun de ces exemples (et dans bien d'autres) la VA est toujours répartie de manière *approximativement poissonienne*, suivant la loi de binomiale avec paramètre n grand et paramètre p petit (même si on ne connais pas la 'vraie' valeur n)

Exemple

Exemple 4.4 Admettons que le nombre d'erreurs par page dans un livre suive une distribution de Poisson avec paramètre $\lambda = \frac{1}{2}$. Calculer la probabilité qu'il y ait au moins une erreur sur la page 27 ...

Autre exemple

Exemple 4.5 On considère l'expérience qui consiste à mesurer le nombre de particules α émises en une seconde par un gramme de matière radioactive. Des expériences ont montré qu'en moyenne le nombre de particules α émises est 3.2.
Donner une bonne approximation de la probabilité qu'au plus deux particules α seront enregistrées ...

Solution Représentons le gramme de matière radioactive comme une collection de n atomes (n est grand). Chacun peut se désintégrer, ceci avec une probabilité de $3.2/n$ pour la durée de mesure et donner un particule α.

On peut alors dire que $X =$ nombre de particules α émises sera approximativement une VA de Poisson de paramètre $\lambda =$ ____

Alors $P(X \leq 2) =$ ...

Espérance et variance de VAs de Poisson : **Intuition**

- Rappelons qu'une VA de Poisson est une approximation d'une VA binomiale de paramètres n et p lorsque n est grand, p est petit et $\lambda = np$
- Soit $X \sim Bin(n, p)$ (n grand, p petit, np modéré) :
  - $E[X] = np = \boldsymbol{\lambda}$
  - $Var(X) = np(1-p) = \lambda(1-p) \approx \boldsymbol{\lambda}$ (si p est petit)
- Il semblerait donc que l'espérance et la variance d'une VA de Poisson soient *toutes les deux égales au paramètre* λ
- Nous pouvons *vérifier cette intuition* par le calcul (qui **ne sera pas examiné**)
- **[si intéressant pour vous, voir ces calculs sur moodle]**

Espérancer d'une VA de Poisson : $\boxed{\textbf{Calcul}}$

$$
\begin{aligned}
E[X] &= \sum_{i=0}^{\infty} \frac{i e^{-\lambda} \lambda^i}{i!} && \underline{\hspace{8cm}} \\
&= \lambda \sum_{i=1}^{\infty} \frac{e^{-\lambda} \lambda^{i-1}}{(i-1)!} && \underline{\hspace{8cm}} \\
&= \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} && (\text{posing } j = i - 1) \\
&= \lambda && \left(\text{since } \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \underline{\hspace{1cm}}\right)
\end{aligned}
$$

Variance of Poisson RVs : redbf $\boxed{\text{Calcul}}$

$$
\begin{aligned}
E[X^2] &= \sum_{i=0}^{\infty} \frac{i^2 e^{-\lambda}\lambda^i}{i!} \\
&= \lambda \sum_{i=1}^{\infty} \frac{i e^{-\lambda}\lambda^{i-1}}{(i-1)!} \\
&= \lambda \sum_{j=0}^{\infty} \frac{(j+1) e^{-\lambda}\lambda^{j}}{j!} \qquad (\text{posing } j = i-1) \\
&= \lambda \left[\sum_{j=0}^{\infty} \frac{j e^{-\lambda}\lambda^{j}}{j!} + \sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^{j}}{j!} \right] \\
&= \lambda\,(\lambda + 1) \qquad [\text{substitution } E(X), \sum_i p(i) = 1]
\end{aligned}
$$

Exemple

Exemple 4.6 Dans une expérience de mutagenèse microbienne, une plaque de bactéries est exposée à un composé, et le nombre de mutants est compté après incubation. Supposons que dans l'expérience d'un composé particulier, le nombre des mutants a une distribution de Poisson avec $\lambda = 9$.

Calculer la probabilité qu'une expérience du composé produise :

(a) 0 mutants

(b) plus que 3 mutants

Autre exemple

Exemple 4.7 Soit Y le nombre de tests avec mutants dans 5 essais indépendants du composé de l'Exemple 4.6.

(a) Quel modèle de probabilité est raisonnable pour Y *??* Expliquer.

(b) Quelles sont les valeurs des paramètres du modèle *??*

(c) Quelle est la probabilité qu'il y ait au moins 1 mutant dans (exactement) 2 des 5 expériences *??*