LOI BINOMIALE ET PRISE DE DÉCISION EN PREMIÈRE

Réflexion sur les enjeux et la mise en œuvre

Stage IREM *Probabilités & Statistique*,
Besançon, 16 février 2012

Yves DUCEL, Bruno SAUSSEREAU
Institut de recherche sur l'enseignement des mathématiques
Université de Franche-Comté

Travail IREM - Université de Franche-Comté

- **Yves DUCEL** et **Bruno SAUSSEREAU**, enseignants-chercheurs à l'Université de Franche-Comté.
- Cet exposé est le fruit d'une réflexion menée dans le cadre du groupe de travail *Probabilités et statistique* de l'IREM de Franche-Comté.
- Large exploitation du travail effectué dans le groupe de rédaction du documents ressource "Probab-stat" de Première.
- Le diaporama sera disponible.

Introduction

Les enjeux

Enjeux culturels

- Répondre à un besoin de la culture du citoyen d'aujourd'hui ;
- Développer une culture de la statistique minimale pour un citoyen ;
- Forger un mode de pensée spécifique pour appréhender l'incertain et le formaliser ;
- Illustrer la pertinence des mathématiques pour répondre à certaines problématiques concrètes (sondages) ;
- Montrer l'interaction des différentes branches des mathématiques entre elles.

Enjeux pour la Première

- Prise en compte du réel dans le cours de mathématique ;
- Mise en place d'un modèle performant et fondamental en statistique ;
- Généralisation de la notion d'intervalle de fluctuation vue en Seconde ;
- Extension de la démarche de prise de décision vue en Seconde ;
- Favoriser une démarche interdisciplinaire au lycée.

La mise en œuvre

La mise en œuvre du programme

- Le propos ne prétend pas statuer sur la mise en œuvre dans la classe,
- il vise plutôt à illustrer les enjeux pointés et à proposer une "philosophie" pour cette mise en œuvre.
- Le discours n'a pas été conçu comme une progression d'enseignement pour la classe,
- il ne colle pas exactement aux attendus du programme (unilatéral, vocabulaire, démarche) et tous les aspects du programme ne sont pas abordés.
- Le document ressource donne une autre approche de certains points développés ci-après.

La problématique

Woburn : énoncé

- Une petite ville des États-Unis, Woburn, a connu 9 cas de leucémie parmi les 5969 garçons de moins de 15 ans sur la période 1969-1979. La fréquence des leucémies pour cette tranche d'âge et pour cette même période aux États-Unis est égale à $0,00052$ (source : *Massachussetts Department of Public Health*). Les autorités concluent qu'il n'y a rien d'étrange dans cette ville. Qu'en pensez-vous ?
(Extrait de *Ressources Baccalauréat professionnel*, page 8)
- Le nombre de cas observés est-il **significatif** d'une situation anormale ou bien est-il simplement dû au hasard ?

Woburn : énoncé

- Une petite ville des États-Unis, Woburn, a connu 9 cas de leucémie parmi les 5969 garçons de moins de 15 ans sur la période 1969-1979. La fréquence des leucémies pour cette tranche d'âge et pour cette même période aux États-Unis est égale à $0,00052$ (source : *Massachussetts Department of Public Health*). Les autorités concluent qu'il n'y a rien d'étrange dans cette ville. Qu'en pensez-vous ?
  (Extrait de *Ressources Baccalauréat professionnel*, page 8)

- Le nombre de cas observés est-il **significatif** d'une situation anormale ou bien est-il simplement dû au hasard ?

Woburn : schéma d'urne

- Dans la suite, la population des États-Unis sera assimilée à une urne contenant 100 000 000 boules rouges ou vertes :
  - les boules rouges, au nombre de 52 000, représentent les personnes atteintes de leucémie,
  - les boules vertes représentent les personnes non atteintes.
- La population des 5969 garçons de moins de 15 ans sur la période 1969-1979 à Woburn sera assimilée à l'observation d'un échantillon de 5969 boules, prélevées au hasard et avec remise dans l'urne.
- Remarque sur le modèle choisi de tirage avec remise.

Convention et notations

Soit un entier $n \geq 1$ et une expérience aléatoire $\mathcal{E}$:

- On notera $\mathcal{E}_n$ l'expérience aléatoire définie par le protocole suivant : *On effectue n épreuves successives de l'expérience $\mathcal{E}$ et on note, en respectant l'ordre, les issues de ces n épreuves successives.*
- On dira pour simplifier que $\mathcal{E}$ est l'expérience **de base** (ou **initiale** ou **de départ**), et que l'expérience $\mathcal{E}_n$ est obtenue par n "répétitions" **à l'identique** de l'expérience de départ $\mathcal{E}$.

Woburn : cadre probabiliste

- L'expérience aléatoire de départ, $\mathcal{E}$: "*Extraire une boule de l'urne et noter sa couleur*", est une expérience de Bernoulli avec $p = 0,00052$.
- L'expérience étudiée est l'expérience $\mathcal{E}_{5969}$ obtenue par 5969 "répétitions" à l'identique de $\mathcal{E}$.
- On associe à l'expérience aléatoire $\mathcal{E}_{5969}$ un modèle probabiliste $(\Omega, \mathbb{P})$ et la variable aléatoire X qui, à toute issue ω de l'expérience $\mathcal{E}_{5969}$ fait correspondre $X(\omega)$:=*nombre de boules rouges comptées dans l'issue* ω.

Woburn : cadre probabiliste

- L'expérience aléatoire de départ, $\mathcal{E}$: "*Extraire une boule de l'urne et noter sa couleur*", est une expérience de Bernoulli avec $p = 0,00052$.
- L'expérience étudiée est l'expérience $\mathcal{E}_{5969}$ obtenue par 5969 "répétitions" à l'identique de $\mathcal{E}$.
- On associe à l'expérience aléatoire $\mathcal{E}_{5969}$ un modèle probabiliste $(\Omega, \mathbb{P})$ et la variable aléatoire X qui, à toute issue ω de l'expérience $\mathcal{E}_{5969}$ fait correspondre $X(\omega) :=$*nombre de boules rouges comptées dans l'issue* ω.

Woburn : cadre probabiliste

- L'expérience aléatoire de départ, $\mathcal{E}$: "*Extraire une boule de l'urne et noter sa couleur*", est une expérience de Bernoulli avec $p = 0,00052$.
- L'expérience étudiée est l'expérience $\mathcal{E}_{5969}$ obtenue par 5969 "répétitions" à l'identique de $\mathcal{E}$.
- On associe à l'expérience aléatoire $\mathcal{E}_{5969}$ un modèle probabiliste $(\Omega, \mathbb{P})$ et la variable aléatoire X qui, à toute issue ω de l'expérience $\mathcal{E}_{5969}$ fait correspondre $X(\omega) :=$*nombre de boules rouges comptées dans l'issue* ω.

Woburn : Étude par simulation

- Observation d'une réalisation de l'expérience $\mathcal{E}_{5969}$.
- Déterminer la probabilité, $\mathbb{P}(X = 9)$, d'obtenir 9 boules rouges lorsqu'on effectue 5969 épreuves de l'expérience "*Extraire une boule de l'urne et noter sa couleur*".
- Estimation par simulation informatique de $\mathbb{P}(X = 9)$ et décision par rapport à la question posée :

Simuler Woburn pour prise de décision

Woburn : Étude par simulation

- Observation d'une réalisation de l'expérience $\mathcal{E}_{5969}$.
- Déterminer la probabilité, $\mathbb{P}(X = 9)$, d'obtenir 9 boules rouges lorsqu'on effectue 5969 épreuves de l'expérience "*Extraire une boule de l'urne et noter sa couleur*".
- Estimation par simulation informatique de $\mathbb{P}(X = 9)$ et décision par rapport à la question posée :

Simuler Woburn pour prise de décision

Woburn : Étude par simulation

- Observation d'une réalisation de l'expérience $\mathcal{E}_{5969}$.
- Déterminer la probabilité, $\mathbb{P}(X = 9)$, d'obtenir 9 boules rouges lorsqu'on effectue 5969 épreuves de l'expérience "*Extraire une boule de l'urne et noter sa couleur*".
- Estimation par simulation informatique de $\mathbb{P}(X = 9)$ et décision par rapport à la question posée :

Simuler Woburn pour prise de décision

Woburn : Étude par simulation

- Observation d'une réalisation de l'expérience $\mathcal{E}_{5969}$.
- Déterminer la probabilité, $\mathbb{P}(X = 9)$, d'obtenir 9 boules rouges lorsqu'on effectue 5969 épreuves de l'expérience "*Extraire une boule de l'urne et noter sa couleur*".
- Estimation par simulation informatique de $\mathbb{P}(X = 9)$ et décision par rapport à la question posée :

Simuler Woburn pour prise de décision

Démarches possibles

- La probabilité $\mathbb{P}(X = 9)$ est trop faible, comparée à celle des valeurs plus probables de X.
- La fréquence observée $f = \dfrac{9}{5969} \simeq 0.001$ est trop grande devant la proportion $p = 0,00052$.
- On considère qu'il y a une **différence significative** entre la réalité observée et ce que l'on attendrait si la situation était normale : on décide que la situation est "anormale".

Commentaires

- Que signifie trop faible ? trop grand ?
- Comment prendre en compte les autres probabilités dans la démarche ?
- Nous avons utilisé une démarche intuitive dans la prise de décision.

Critique

- La démarche utilisée n'est pas formalisée ni quantifiée.
- La probabilité est estimée, mais on ne connaît pas son expression mathématique exacte.
- Nous prenons une décision, mais nous ne connaissons pas le risque de nous tromper en la prenant.
- La méthode vue en Seconde permettrait d'évaluer ce risque d'erreur, mais elle n'est pas utilisée.
- Les conditions de validité de la méthode vue en Seconde ne sont pas réalisées, car p très petit.

Objectifs

Nous nous proposons maintenant de :

- Étudier le modèle probabiliste associé au schéma d'urne de l'expérience $\mathcal{E}_{5969}$.
- Préciser la connaissance mathématique de la loi de la variable X introduite.
- Formaliser une méthode de prise de décision :
  - sur la base de la démarche initiée en Seconde,
  - valable sans restriction pour tout réel $0 < p < 1$, tout entier $n \geq 1$.
  - qui permette de retrouver et justifier mathématiquement *a posteriori* les résultats vus en Seconde.

Objectifs

Nous nous proposons maintenant de :

- Étudier le modèle probabiliste associé au schéma d'urne de l'expérience $\mathcal{E}_{5969}$.
- Préciser la connaissance mathématique de la loi de la variable X introduite.
- Formaliser une méthode de prise de décision :
  - sur la base de la démarche initiée en Seconde,
  - valable sans restriction pour tout réel $0 < p < 1$, tout entier $n \geq 1$.
  - qui permette de retrouver et justifier mathématiquement *a posteriori* les résultats vus en Seconde.

Objectifs

Nous nous proposons maintenant de :

- Étudier le modèle probabiliste associé au schéma d'urne de l'expérience $\mathcal{E}_{5969}$.
- Préciser la connaissance mathématique de la loi de la variable X introduite.
- Formaliser une méthode de prise de décision :
  - sur la base de la démarche initiée en Seconde,
  - valable sans restriction pour tout réel $0 < p < 1$, tout entier $n \geq 1$.
  - qui permette de retrouver et justifier mathématiquement *a posteriori* les résultats vus en Seconde.

Objectifs

Nous nous proposons maintenant de :

- Étudier le modèle probabiliste associé au schéma d'urne de l'expérience $\mathcal{E}_{5969}$.
- Préciser la connaissance mathématique de la loi de la variable X introduite.
- Formaliser une méthode de prise de décision :
  - sur la base de la démarche initiée en Seconde,
  - valable sans restriction pour tout réel $0 < p < 1$, tout entier $n \geq 1$.
  - qui permette de retrouver et justifier mathématiquement *a posteriori* les résultats vus en Seconde.

Objectifs

Nous nous proposons maintenant de :

- Étudier le modèle probabiliste associé au schéma d'urne de l'expérience $\mathcal{E}_{5969}$.
- Préciser la connaissance mathématique de la loi de la variable X introduite.
- Formaliser une méthode de prise de décision :
  - sur la base de la démarche initiée en Seconde,
  - valable sans restriction pour tout réel $0 < p < 1$, tout entier $n \geq 1$.
  - qui permette de retrouver et justifier mathématiquement *a posteriori* les résultats vus en Seconde.

Objectifs

Nous nous proposons maintenant de :

- Étudier le modèle probabiliste associé au schéma d'urne de l'expérience $\mathcal{E}_{5969}$.
- Préciser la connaissance mathématique de la loi de la variable X introduite.
- Formaliser une méthode de prise de décision :
  - sur la base de la démarche initiée en Seconde,
  - valable sans restriction pour tout réel $0 < p < 1$, tout entier $n \geq 1$.
  - qui permette de retrouver et justifier mathématiquement *a posteriori* les résultats vus en Seconde.

Partie I

La loi binomiale

Modèles et arbres pondérés

Convention et notations

Soit un entier $n \geq 1$ et une expérience aléatoire $\mathcal{E}$:

- On notera $\mathcal{E}_n$ l'expérience aléatoire définie par le protocole suivant : *On effectue n épreuves successives de l'expérience $\mathcal{E}$ et on note, en respectant l'ordre, les issues de ces n épreuves successives.*
- On dira pour simplifier que $\mathcal{E}$ est l'expérience **de base** (ou **initiale** ou **de départ**), et que l'expérience $\mathcal{E}_n$ est obtenue par n "répétitions" **à l'identique** de l'expérience de départ $\mathcal{E}$.

Protocole $\mathcal{E}$

- Pour simplifier le raisonnement, considérons l'expérience de base $\mathcal{E}$ suivante :

  Soit une urne contenant 2 boules vertes et 3 boules rouges, indistinguables au toucher mais identifiables à la vue (boules numérotées par exemple). On extrait une boule de l'urne et on note sa couleur.

- Nous allons étudier deux modèles associés à l'expérience aléatoire $\mathcal{E}_3$ qui consiste à répéter trois fois à l'identique l'expérience de base.

Le modèle probabiliste $(\Omega, \mathbb{P})$ de $\mathcal{E}_3$

- Choix du modèle probabiliste $(\Omega, \mathbb{P})$: Quel Ω ? Quel $\mathbb{P}$?
- Quel Ω ? Identifier toutes les issues possibles
  - * On prend pour univers des possibles l'ensemble des triplets formés avec les deux couleurs r ou v :
    $\Omega := \{rrr, rrv, rvr, rvv, vrr, vrv, vvr, vvv\}$.
- Choix de la probabilité $\mathbb{P}$?
  - * L'équiprobabilité ne convient pas : principe de raison insuffisante en défaut.
  - * Intérêt de passer par un modèle intermédiaire $(\Omega', \mathbb{P}')$ pour décrire $\mathcal{E}_3$.

Le modèle probabiliste $(\Omega, \mathbb{P})$ de $\mathcal{E}_3$

- Choix du modèle probabiliste $(\Omega, \mathbb{P})$: Quel Ω ? Quel $\mathbb{P}$?
- Quel Ω ? Identifier toutes les issues possibles
  - On prend pour univers des possibles l'ensemble des triplets formés avec les deux couleurs r ou v : $\Omega := \{rrr,\ rrv,\ rvr,\ rvv,\ vrr,\ vrv,\ vvr,\ vvv\}$.
- Choix de la probabilité $\mathbb{P}$?
  - L'équiprobabilité ne convient pas : principe de raison insuffisante en défaut.
  - Intérêt de passer par un modèle intermédiaire $(\Omega', \mathbb{P}')$ pour décrire $\mathcal{E}_3$.

Le modèle probabiliste $(\Omega, \mathbb{P})$ de $\mathcal{E}_3$

- Choix du modèle probabiliste $(\Omega, \mathbb{P})$: Quel Ω ? Quel $\mathbb{P}$?
- Quel Ω ? Identifier toutes les issues possibles
  - * On prend pour univers des possibles l'ensemble des triplets formés avec les deux couleurs r ou v :
    $\Omega := \{rrr,\ rrv,\ rvr,\ rvv,\ vrr,\ vrv,\ vvr,\ vvv\}$.
- Choix de la probabilité $\mathbb{P}$?
  - * L'équiprobabilité ne convient pas : principe de raison insuffisante en défaut.
  - * Intérêt de passer par un modèle intermédiaire $(\Omega', \mathbb{P}')$ pour décrire $\mathcal{E}_3$.

Le modèle probabiliste $(\Omega, \mathbb{P})$ de $\mathcal{E}_3$

- Choix du modèle probabiliste $(\Omega, \mathbb{P})$: Quel Ω ? Quel $\mathbb{P}$?
- Quel Ω ? Identifier toutes les issues possibles
  * On prend pour univers des possibles l'ensemble des triplets formés avec les deux couleurs r ou v :
    $\Omega := \{rrr,\ rrv,\ rvr,\ rvv,\ vrr,\ vrv,\ vvr,\ vvv\}$.
- Choix de la probabilité $\mathbb{P}$?
  * L'équiprobabilité ne convient pas : principe de raison insuffisante en défaut.
  * Intérêt de passer par un modèle intermédiaire $(\Omega', \mathbb{P}')$ pour décrire $\mathcal{E}_3$.

Le modèle probabiliste $(\Omega, \mathbb{P})$ de $\mathcal{E}_3$

- Choix du modèle probabiliste $(\Omega, \mathbb{P})$: Quel Ω ? Quel $\mathbb{P}$?
- Quel Ω ? Identifier toutes les issues possibles
  - * On prend pour univers des possibles l'ensemble des triplets formés avec les deux couleurs r ou v :
    $\Omega := \{rrr,\ rrv,\ rvr,\ rvv,\ vrr,\ vrv,\ vvr,\ vvv\}$.
- Choix de la probabilité $\mathbb{P}$?
  - * L'équiprobabilité ne convient pas : principe de raison insuffisante en défaut.
  - * Intérêt de passer par un modèle intermédiaire $(\Omega', \mathbb{P}')$ pour décrire $\mathcal{E}_3$.

Le modèle probabiliste $(\Omega, \mathbb{P})$ de $\mathcal{E}_3$

- Choix du modèle probabiliste $(\Omega, \mathbb{P})$: Quel Ω ? Quel $\mathbb{P}$?
- Quel Ω ? Identifier toutes les issues possibles
  - On prend pour univers des possibles l'ensemble des triplets formés avec les deux couleurs r ou v :
    $\Omega := \{rrr,\ rrv,\ rvr,\ rvv,\ vrr,\ vrv,\ vvr,\ vvv\}$.
- Choix de la probabilité $\mathbb{P}$?
  - L'équiprobabilité ne convient pas : principe de raison insuffisante en défaut.
  - Intérêt de passer par un modèle intermédiaire $(\Omega', \mathbb{P}')$ pour décrire $\mathcal{E}_3$.

Le modèle intermédiaire $(\Omega', \mathbb{P}')$ de $\mathcal{E}_3$

On choisit le modèle probabiliste $(\Omega', \mathbb{P}')$ de façon à prendre en compte la boule elle-même, et pas seulement sa couleur (on peut imaginer que les boules sont numérotées et qu'on note sa couleur et son numéro).

- Quel Ω' ? Identifier toutes les issues possibles
  - Ω' est l'ensemble des triplets de boules. Il y a 125 issues possibles.
- Choix de la probabilité $\mathbb{P}'$?
  - L'équiprobabilité convient : principe de raison insuffisante.
  - On prend pour $\mathbb{P}'$ l'équiprobabilité sur Ω'.

♣ Remarque sur les arbres pondérés : Justifier la pondération de l'arbre

Le modèle intermédiaire $(\Omega', \mathbb{P}')$ de $\mathcal{E}_3$

On choisit le modèle probabiliste $(\Omega', \mathbb{P}')$ de façon à prendre en compte la boule elle-même, et pas seulement sa couleur (on peut imaginer que les boules sont numérotées et qu'on note sa couleur et son numéro).

- Quel Ω' ? Identifier toutes les issues possibles
  - Ω' est l'ensemble des triplets de boules. Il y a 125 issues possibles.
- Choix de la probabilité $\mathbb{P}'$?
  - L'équiprobabilité convient : principe de raison insuffisante.
  - On prend pour $\mathbb{P}'$ l'équiprobabilité sur Ω'.

♣ Remarque sur les arbres pondérés : Justifier la pondération de l'arbre

Le modèle intermédiaire $(\Omega', \mathbb{P}')$ de $\mathcal{E}_3$

On choisit le modèle probabiliste $(\Omega', \mathbb{P}')$ de façon à prendre en compte la boule elle-même, et pas seulement sa couleur (on peut imaginer que les boules sont numérotées et qu'on note sa couleur et son numéro).

- Quel Ω' ? Identifier toutes les issues possibles
  - Ω' est l'ensemble des triplets de boules. Il y a 125 issues possibles.
- Choix de la probabilité $\mathbb{P}'$?
  - L'équiprobabilité convient : principe de raison insuffisante.
  - On prend pour $\mathbb{P}'$ l'équiprobabilité sur Ω'.

♣ Remarque sur les arbres pondérés : Justifier la pondération de l'arbre

Le modèle intermédiaire $(\Omega', \mathbb{P}')$ de $\mathcal{E}_3$

On choisit le modèle probabiliste $(\Omega', \mathbb{P}')$ de façon à prendre en compte la boule elle-même, et pas seulement sa couleur (on peut imaginer que les boules sont numérotées et qu'on note sa couleur et son numéro).

- Quel Ω' ? Identifier toutes les issues possibles
  - Ω' est l'ensemble des triplets de boules. Il y a 125 issues possibles.
- Choix de la probabilité $\mathbb{P}'$?
  - L'équiprobabilité convient : principe de raison insuffisante.
  - On prend pour $\mathbb{P}'$ l'équiprobabilité sur Ω'.

♣ Remarque sur les arbres pondérés : Justifier la pondération de l'arbre

Le modèle intermédiaire $(\Omega', \mathbb{P}')$ de $\mathcal{E}_3$

On choisit le modèle probabiliste $(\Omega', \mathbb{P}')$ de façon à prendre en compte la boule elle-même, et pas seulement sa couleur (on peut imaginer que les boules sont numérotées et qu'on note sa couleur et son numéro).

- Quel Ω' ? Identifier toutes les issues possibles
  - * Ω' est l'ensemble des triplets de boules. Il y a 125 issues possibles.
- Choix de la probabilité $\mathbb{P}'$?
  - * L'équiprobabilité convient : principe de raison insuffisante.
  - * On prend pour $\mathbb{P}'$ l'équiprobabilité sur Ω'.

♣ Remarque sur les arbres pondérés : Justifier la pondération de l'arbre

Le modèle intermédiaire $(\Omega', \mathbb{P}')$ de $\mathcal{E}_3$

On choisit le modèle probabiliste $(\Omega', \mathbb{P}')$ de façon à prendre en compte la boule elle-même, et pas seulement sa couleur (on peut imaginer que les boules sont numérotées et qu'on note sa couleur et son numéro).

- Quel Ω' ? Identifier toutes les issues possibles
  - * Ω' est l'ensemble des triplets de boules. Il y a 125 issues possibles.
- Choix de la probabilité $\mathbb{P}'$?
  - * L'équiprobabilité convient : principe de raison insuffisante.
  - * On prend pour $\mathbb{P}'$ l'équiprobabilité sur Ω'.

♣ Remarque sur les arbres pondérés : Justifier la pondération de l'arbre

Remarques sur les modèles associés à $\mathcal{E}_3$

- Tous les événements de $(\Omega, \mathbb{P})$ peuvent être modélisés avec $(\Omega', \mathbb{P}')$. Mais la réciproque est fausse.
- Le modèle $(\Omega', \mathbb{P}')$ est plus riche en événements que le modèle $(\Omega, \mathbb{P})$.
- Les événements élémentaires pour $(\Omega, \mathbb{P})$ ne sont pas modélisés par des événements élémentaires pour $(\Omega', \mathbb{P}')$.
- On prend pour probabilité d'un événement élémentaire de $(\Omega, \mathbb{P})$, la probabilité de l'événement correspondant modélisé avec $(\Omega', \mathbb{P}')$.
- Le modèle intermédiaire d'équiprobabilité $(\Omega', \mathbb{P}')$ nous permet de justifier la règle du produit des probabilités le long d'une branche de l'arbre pondéré associé à $(\Omega, \mathbb{P})$.

Conclusion sur $(\Omega, \mathbb{P})$

Le modèle probabiliste $(\Omega, \mathbb{P})$ retenu pour décrire l'expérience à trois répétitions $\mathcal{E}_3$ peut être résumé par le tableau :

ω	rrr	rrv	rvr	rvv	vrr	vrv	vvr	vvv
$\mathbb{P}(\{\omega\})$	$\frac{8}{125}$	$\frac{12}{125}$	$\frac{12}{125}$	$\frac{18}{125}$	$\frac{12}{125}$	$\frac{18}{125}$	$\frac{18}{125}$	$\frac{27}{125}$

Généralisation à n répétitions et paramètre p

- Expérience de départ $\mathcal{E}$ de paramètre p. Le modèle probabiliste $(\Omega, \mathbb{P})$ se généralise au cas de l'expérience à n répétitions $\mathcal{E}_n$:
  - Ω est l'ensemble des n-uplets formés avec les deux couleurs r et v.
  - $\mathbb{P}$ est définie par ses valeurs pour chacun des événements élémentaires : la probabilité d'obtenir une issue contenant exactement k fois la couleur r est égale à $p^k q^{n-k}$, où $q := 1 - p$.
- Le nombre d'issues contenant exactement k fois la couleur r est égal au nombre de branches de l'arbre pondéré contenant exactement k boules rouges.

♣ Notation : On notera $\binom{n}{k}$ ce nombre. Lire « *binomial n, k* ».

Généralisation à n répétitions et paramètre p

- Expérience de départ $\mathcal{E}$ de paramètre p. Le modèle probabiliste $(\Omega, \mathbb{P})$ se généralise au cas de l'expérience à n répétitions $\mathcal{E}_n$:
  - Ω est l'ensemble des n-uplets formés avec les deux couleurs r et v.
  - $\mathbb{P}$ est définie par ses valeurs pour chacun des événements élémentaires : la probabilité d'obtenir une issue contenant exactement k fois la couleur r est égale à $p^k q^{n-k}$, où $q := 1 - p$.
- Le nombre d'issues contenant exactement k fois la couleur r est égal au nombre de branches de l'arbre pondéré contenant exactement k boules rouges.

♣ Notation : On notera $\binom{n}{k}$ ce nombre. Lire « *binomial n, k* ».

Généralisation à n répétitions et paramètre p

- Expérience de départ $\mathcal{E}$ de paramètre p. Le modèle probabiliste $(\Omega, \mathbb{P})$ se généralise au cas de l'expérience à n répétitions $\mathcal{E}_n$:
  - Ω est l'ensemble des n-uplets formés avec les deux couleurs r et v.
  - $\mathbb{P}$ est définie par ses valeurs pour chacun des événements élémentaires : la probabilité d'obtenir une issue contenant exactement k fois la couleur r est égale à $p^k q^{n-k}$, où $q := 1 - p$.
- Le nombre d'issues contenant exactement k fois la couleur r est égal au nombre de branches de l'arbre pondéré contenant exactement k boules rouges.

♣ Notation : On notera $\binom{n}{k}$ ce nombre. Lire «*binomial n, k*».

Généralisation à n répétitions et paramètre p

- Expérience de départ $\mathcal{E}$ de paramètre p. Le modèle probabiliste $(\Omega, \mathbb{P})$ se généralise au cas de l'expérience à n répétitions $\mathcal{E}_n$:
  - Ω est l'ensemble des n-uplets formés avec les deux couleurs r et v.
  - $\mathbb{P}$ est définie par ses valeurs pour chacun des événements élémentaires : la probabilité d'obtenir une issue contenant exactement k fois la couleur r est égale à $p^k q^{n-k}$, où $q := 1 - p$.
- Le nombre d'issues contenant exactement k fois la couleur r est égal au nombre de branches de l'arbre pondéré contenant exactement k boules rouges.

♣ Notation : On notera $\binom{n}{k}$ ce nombre. Lire «*binomial n, k*».

Généralisation à n répétitions et paramètre p

- Expérience de départ $\mathcal{E}$ de paramètre p. Le modèle probabiliste $(\Omega, \mathbb{P})$ se généralise au cas de l'expérience à n répétitions $\mathcal{E}_n$:
  - Ω est l'ensemble des n-uplets formés avec les deux couleurs r et v.
  - $\mathbb{P}$ est définie par ses valeurs pour chacun des événements élémentaires : la probabilité d'obtenir une issue contenant exactement k fois la couleur r est égale à $p^k q^{n-k}$, où $q := 1 - p$.
- Le nombre d'issues contenant exactement k fois la couleur r est égal au nombre de branches de l'arbre pondéré contenant exactement k boules rouges.

♣ Notation : On notera $\binom{n}{k}$ ce nombre. Lire «*binomial* n, k».

Variables aléatoires

Exemples de variables aléatoires sur $(\Omega, \mathbb{P})$:

- La variable X **binomiale de paramètres** $n = 3$ **et** $p = 2/5$:

$$\omega \in \Omega \mapsto X(\omega) = \text{ nombre de boules rouges dans l'issue } \omega.$$

- La variable Y **géométrique tronquée de paramètres** $n = 3$ **et** $p = 2/5$:

$$\omega \in \Omega \mapsto Y(\omega)$$

définie de la façon suivante :

  - * s'il y a au moins une boule rouge dans l'issue ω, $Y(\omega)$ est alors le rang de la première boule rouge ;
  - * s'il n'y a aucune boule rouge dans l'issue ω, on pose $Y(\omega) = 0$.

Variables aléatoires

Exemples de variables aléatoires sur $(\Omega, \mathbb{P})$:

- La variable X **binomiale de paramètres** $n = 3$ **et** $p = 2/5$:

$$\omega \in \Omega \mapsto X(\omega) = \text{ nombre de boules rouges dans l'issue } \omega.$$

- La variable Y **géométrique tronquée de paramètres** $n = 3$ **et** $p = 2/5$:

$$\omega \in \Omega \mapsto Y(\omega)$$

  définie de la façon suivante :
  - s'il y a au moins une boule rouge dans l'issue ω, $Y(\omega)$ est alors le rang de la première boule rouge ;
  - s'il n'y a aucune boule rouge dans l'issue ω, on pose $Y(\omega) = 0$.

Variables aléatoires

Exemples de variables aléatoires sur $(\Omega, \mathbb{P})$:

- La variable X **binomiale de paramètres** $n = 3$ **et** $p = 2/5$:

  $$\omega \in \Omega \mapsto X(\omega) = \text{ nombre de boules rouges dans l'issue } \omega.$$

- La variable Y **géométrique tronquée de paramètres** $n = 3$ **et** $p = 2/5$:

  $$\omega \in \Omega \mapsto Y(\omega)$$

  définie de la façon suivante :

  * s'il y a au moins une boule rouge dans l'issue ω, $Y(\omega)$ est alors le rang de la première boule rouge ;
  * s'il n'y a aucune boule rouge dans l'issue ω, on pose $Y(\omega) = 0$.

Variables aléatoires

Exemples de variables aléatoires sur $(\Omega, \mathbb{P})$:

- La variable X **binomiale de paramètres** $n = 3$ **et** $p = 2/5$:

  $$\omega \in \Omega \mapsto X(\omega) = \text{ nombre de boules rouges dans l'issue } \omega.$$

- La variable Y **géométrique tronquée de paramètres** $n = 3$ **et** $p = 2/5$:

  $$\omega \in \Omega \mapsto Y(\omega)$$

  définie de la façon suivante :
  * s'il y a au moins une boule rouge dans l'issue ω, $Y(\omega)$ est alors le rang de la première boule rouge ;
  * s'il n'y a aucune boule rouge dans l'issue ω, on pose $Y(\omega) = 0$.

Découverte de la loi binomiale

Loi de la variable X

- La variable binomiale X prend ses valeurs dans $\{0,\ 1,\ 2,\ 3\}$.
- Loi de la v.a. X : Déterminer la loi
- La loi de X peut être résumée par le tableau :

x	0	1	2	3
$\{X = x\}$	$\{vvv\}$	$\{rvv, vrv, vvr\}$	$\{rrv, vrr, rvr\}$	$\{rrr\}$
$\mathbb{P}(X = x)$	$\left(\frac{3}{5}\right)^3$	$3\left(\frac{2}{5}\right)\left(\frac{3}{5}\right)^2$	$3\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)$	$\left(\frac{2}{5}\right)^3$

Loi de la variable X

- La variable binomiale X prend ses valeurs dans $\{0,\ 1,\ 2,\ 3\}$.
- Loi de la v.a. X : Déterminer la loi
- La loi de X peut être résumée par le tableau :

x	0	1	2	3
$\{X = x\}$	$\{vvv\}$	$\{rvv, vrv, vvr\}$	$\{rrv, vrr, rvr\}$	$\{rrr\}$
$\mathbb{P}(X = x)$	$\left(\frac{3}{5}\right)^3$	$3\left(\frac{2}{5}\right)\left(\frac{3}{5}\right)^2$	$3\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)$	$\left(\frac{2}{5}\right)^3$

Loi de la variable X

- La variable binomiale X prend ses valeurs dans $\{0,\ 1,\ 2,\ 3\}$.
- Loi de la v.a. X : Déterminer la loi
- La loi de X peut être résumée par le tableau :

x	0	1	2	3
$\{X = x\}$	$\{vvv\}$	$\{rvv, vrv, vvr\}$	$\{rrv, vrr, rvr\}$	$\{rrr\}$
$\mathbb{P}(X = x)$	$\left(\frac{3}{5}\right)^3$	$3\left(\frac{2}{5}\right)\left(\frac{3}{5}\right)^2$	$3\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)$	$\left(\frac{2}{5}\right)^3$

Généralisation à n répétitions et paramètre p

Expérience de départ $\mathcal{E}$ de paramètre p avec n répétitions.

- La variable aléatoire X prend ses valeurs dans $\{0,\ 1,\ 2,\ 3,\ \cdots,\ n-1,\ n\}$.
- La loi de la variable aléatoire est donnée, pour tout $k \in \{0,\ 1,\ 2,\ 3,\ \cdots,\ n-1,\ n\}$, par la relation :

$$\mathbb{P}(X=k) = \binom{n}{k} p^k q^{n-k}.$$

Généralisation à n répétitions et paramètre p

Expérience de départ $\mathcal{E}$ de paramètre p avec n répétitions.

- La variable aléatoire X prend ses valeurs dans $\{0,\ 1,\ 2,\ 3,\ \cdots,\ n-1,\ n\}$.
- La loi de la variable aléatoire est donnée, pour tout $k \in \{0,\ 1,\ 2,\ 3,\ \cdots,\ n-1,\ n\}$, par la relation :

$$\mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}.$$

Propriétés et détermination des nombres $\binom{n}{k}$

Quelques relations entre les $\binom{n}{k}$

- $$\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\binom{n}{3}+\cdots+\binom{n}{k}+\cdots+\binom{n}{n-1}+\binom{n}{n}=2^n.$$
- $\binom{n}{0}=\binom{n}{n}=1$, et pour tout entier k tel que $0 \leq k \leq n$,
$$\binom{n}{k}=\binom{n}{n-k}.$$ Justifier la relation
- La **relation de Pascal** : pour tout entier k tel que $0 \leq k \leq n-1$,
$$\binom{n+1}{k+1}=\binom{n}{k+1}+\binom{n}{k}.$$ Justifier la relation

Quelques relations entre les $\binom{n}{k}$

- $$\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\binom{n}{3}+\cdots+\binom{n}{k}+\cdots+\binom{n}{n-1}+\binom{n}{n}=2^n.$$
- $\binom{n}{0}=\binom{n}{n}=1$, et pour tout entier k tel que $0\leq k\leq n$,
$$\binom{n}{k}=\binom{n}{n-k}.$$ Justifier la relation
- **La relation de Pascal** : pour tout entier k tel que $0\leq k\leq n-1$,
$$\binom{n+1}{k+1}=\binom{n}{k+1}+\binom{n}{k}.$$ Justifier la relation

Quelques relations entre les $\binom{n}{k}$

- $$\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\binom{n}{3}+\cdots+\binom{n}{k}+\cdots+\binom{n}{n-1}+\binom{n}{n}=2^n.$$
- $\binom{n}{0}=\binom{n}{n}=1$, et pour tout entier k tel que $0 \leq k \leq n$,
$$\binom{n}{k}=\binom{n}{n-k}.$$ Justifier la relation
- La **relation de Pascal** : pour tout entier k tel que $0 \leq k \leq n-1$,
$$\binom{n+1}{k+1}=\binom{n}{k+1}+\binom{n}{k}.$$ Justifier la relation

Le triangle arithmétique

Conséquence : le **triangle arithmétique** donne un algorithme de calcul des $\binom{n}{k}$:

- Le triangle arithmétique pour $n = 1, 2, 3, 4, 5, 6$ et 7 :

$n = 1$	1	1	0	0	0	0	0	0	0	0
$n = 2$	1	2	1	0	0	0	0	0	0	0
$n = 3$	1	3	3	1	0	0	0	0	0	0
$n = 4$	1	4	6	4	1	0	0	0	0	0
$n = 5$	1	5	10	10	5	1	0	0	0	0
$n = 6$	1	6	15	20	15	6	1	0	0	0
$n = 7$	1	7	21	35	35	21	7	1	0	0

- Histoire : Les triangles arithmétiques de Al-Karaji (Xe siècle) et de Pascal (XVIIe siècle) : Voir les triangles

Le triangle arithmétique

Conséquence : le **triangle arithmétique** donne un algorithme de calcul des $\binom{n}{k}$:

- Le triangle arithmétique pour $n = 1, 2, 3, 4, 5, 6$ et 7 :

$n = 1$	1	1	0	0	0	0	0	0	0	0
$n = 2$	1	2	1	0	0	0	0	0	0	0
$n = 3$	1	3	3	1	0	0	0	0	0	0
$n = 4$	1	4	6	4	1	0	0	0	0	0
$n = 5$	1	5	10	10	5	1	0	0	0	0
$n = 6$	1	6	15	20	15	6	1	0	0	0
$n = 7$	1	7	21	35	35	21	7	1	0	0

- Histoire : Les triangles arithmétiques de Al-Karaji (Xe siècle) et de Pascal (XVIIe siècle) : Voir les triangles

Moments d'une variable binomiale

Espérance de X, $n = 1, 2, 3, 4$: calculs

$$\mathbb{E}(X) = \sum_{k=0}^{k=n} k\mathbb{P}(X = k) = \sum_{k=0}^{k=n} k\binom{n}{k} p^k q^{n-k} \text{ avec } p + q = 1.$$

- Cas $n = 1$: $\mathbb{E}(X) = 0q + 1p = p$.
- Cas $n = 2$: $\mathbb{E}(X) = 2pq + 2p^2 = 2p(p + q) = 2p$.
- Cas $n = 3$: $\mathbb{E}(X) = 3pq^2 + 2 \cdot 3p^2q + 3p^3 = 3p(q^2 + 2pq + p^2) = 3p(p + q)^2 = 3p$.
- Cas $n = 4$: $\mathbb{E}(X) = 4pq^3 + 2 \cdot 6p^2q^2 + 3 \cdot 4p^3q + 4p^4 = 4p(q^3 + 3p^2q + 3pq^2 + p^3) = 4p(p + q)^3 = 4p$.

Espérance de X, $n = 1, 2, 3, 4$: calculs

$$\mathbb{E}(X) = \sum_{k=0}^{k=n} k\mathbb{P}(X = k) = \sum_{k=0}^{k=n} k\binom{n}{k} p^k q^{n-k} \text{ avec } p + q = 1.$$

- Cas $n = 1$: $\mathbb{E}(X) = 0q + 1p = p$.
- Cas $n = 2$: $\mathbb{E}(X) = 2pq + 2p^2 = 2p(p + q) = 2p$.
- Cas $n = 3$: $\mathbb{E}(X) = 3pq^2 + 2 \cdot 3p^2q + 3p^3 = 3p(q^2 + 2pq + p^2) = 3p(p + q)^2 = 3p$.
- Cas $n = 4$: $\mathbb{E}(X) = 4pq^3 + 2 \cdot 6p^2q^2 + 3 \cdot 4p^3q + 4p^4 = 4p(q^3 + 3p^2q + 3pq^2 + p^3) = 4p(p + q)^3 = 4p$.

Espérance de X, $n = 1, 2, 3, 4$: calculs

$$\mathbb{E}(X) = \sum_{k=0}^{k=n} k\mathbb{P}(X = k) = \sum_{k=0}^{k=n} k\binom{n}{k} p^k q^{n-k} \text{ avec } p + q = 1.$$

- Cas $n = 1$: $\mathbb{E}(X) = 0q + 1p = p$.
- Cas $n = 2$: $\mathbb{E}(X) = 2pq + 2p^2 = 2p(p + q) = 2p$.
- Cas $n = 3$: $\mathbb{E}(X) = 3pq^2 + 2 \cdot 3p^2q + 3p^3 = 3p(q^2 + 2pq + p^2) = 3p(p + q)^2 = 3p$.
- Cas $n = 4$: $\mathbb{E}(X) = 4pq^3 + 2 \cdot 6p^2q^2 + 3 \cdot 4p^3q + 4p^4 = 4p(q^3 + 3p^2q + 3pq^2 + p^3) = 4p(p + q)^3 = 4p$.

Espérance de X, $n = 1, 2, 3, 4$: calculs

$$\mathbb{E}(X) = \sum_{k=0}^{k=n} k\mathbb{P}(X = k) = \sum_{k=0}^{k=n} k\binom{n}{k} p^k q^{n-k} \text{ avec } p + q = 1.$$

- Cas $n = 1$: $\mathbb{E}(X) = 0q + 1p = p$.
- Cas $n = 2$: $\mathbb{E}(X) = 2pq + 2p^2 = 2p(p + q) = 2p$.
- Cas $n = 3$: $\mathbb{E}(X) = 3pq^2 + 2 \cdot 3p^2q + 3p^3 = 3p(q^2 + 2pq + p^2) = 3p(p + q)^2 = 3p$.
- Cas $n = 4$: $\mathbb{E}(X) = 4pq^3 + 2 \cdot 6p^2q^2 + 3 \cdot 4p^3q + 4p^4 = 4p(q^3 + 3p^2q + 3pq^2 + p^3) = 4p(p + q)^3 = 4p$.

Moments d'une variable binomiale : cas général

- Généralisation à n quelconque : $\mathbb{E}(X) = np$.
- La variance de X est $\sigma_X^2 = npq$.
- ♣ Au passage ... : Relation entre le triangle arithmétique et les identités remarquables (**Binôme de Newton**) :

$$\binom{n}{0} q^n + \binom{n}{1} p^1 q^{n-1} + \cdots + \binom{n}{k} p^k q^{n-k} + \cdots + \binom{n}{n} p^n = (p+q)^n .$$

- ♣ Vocabulaire : Dans la suite de l'exposé, on parlera de "**grande binomiale**" si $n > 25$ et $0,2 < p < 0,8$, et dans le cas contraire de "**petite binomiale**".

Moments d'une variable binomiale : cas général

- Généralisation à n quelconque : $\mathbb{E}(X) = np$.
- La variance de X est $\sigma_X^2 = npq$.
- Au passage ... : Relation entre le triangle arithmétique et les identités remarquables (**Binôme de Newton**) :

$$\binom{n}{0} q^n + \binom{n}{1} p^1 q^{n-1} + \cdots + \binom{n}{k} p^k q^{n-k} + \cdots + \binom{n}{n} p^n = (p+q)^n .$$

- Vocabulaire : Dans la suite de l'exposé, on parlera de "**grande binomiale**" si $n > 25$ et $0,2 < p < 0,8$, et dans le cas contraire de "**petite binomiale**".

Moments d'une variable binomiale : cas général

- Généralisation à n quelconque : $\mathbb{E}(X) = np$.
- La variance de X est $\sigma_X^2 = npq$.
- ♣ Au passage ... : Relation entre le triangle arithmétique et les identités remarquables (**Binôme de Newton**) :

$$\binom{n}{0} q^n + \binom{n}{1} p^1 q^{n-1} + \cdots + \binom{n}{k} p^k q^{n-k} + \cdots + \binom{n}{n} p^n = (p+q)^n .$$

- ♣ Vocabulaire : Dans la suite de l'exposé, on parlera de "**grande binomiale**" si $n > 25$ et $0,2 < p < 0,8$, et dans le cas contraire de "**petite binomiale**".

Moments d'une variable binomiale : cas général

- Généralisation à n quelconque : $\mathbb{E}(X) = np$.
- La variance de X est $\sigma_X^2 = npq$.

♣ Au passage ... : Relation entre le triangle arithmétique et les identités remarquables (**Binôme de Newton**) :

$$\binom{n}{0} q^n + \binom{n}{1} p^1 q^{n-1} + \cdots + \binom{n}{k} p^k q^{n-k} + \cdots + \binom{n}{n} p^n = (p+q)^n .$$

♣ Vocabulaire : Dans la suite de l'exposé, on parlera de "**grande binomiale**" si $n > 25$ et $0,2 < p < 0,8$, et dans le cas contraire de "**petite binomiale**".

Espérance et variance : conjectures

Autres approches possibles (cf. document ressource *Statistiques et probabilités : classe de Première générale et technologique*)

Conjecturer à partir de graphiques : Graphiques

Loi binomiale et tableur

- Calculs et représentations graphiques : utilisation de la fonction LOI.BINOMIALE des tableurs.
- Influence des paramètres n et p sur : l'allure (symétrie ?) et la position du diagramme, la monotonie, la dispersion de X.

Voir grandes binomiales avec tableur | Voir petites binomiales avec tableur

- Allure symétrique "en cloche" centrée sur np des grandes binomiales.
- Déplacement vers la droite du diagramme à n fixé en fonction de la croissance de p.
- Dispersion maximale en $p = 0,5$: étudier la fonction $p \in [0,1] \mapsto np(1-p)$.

Loi binomiale et tableur

- Calculs et représentations graphiques : utilisation de la fonction LOI.BINOMIALE des tableurs.
- Influence des paramètres n et p sur : l'allure (symétrie ?) et la position du diagramme, la monotonie, la dispersion de X.

Voir grandes binomiales avec tableur | Voir petites binomiales avec tableur

- Allure symétrique "en cloche" centrée sur np des grandes binomiales.
- Déplacement vers la droite du diagramme à n fixé en fonction de la croissance de p.
- Dispersion maximale en $p = 0,5$: étudier la fonction $p \in [0,1] \mapsto np(1-p)$.

Loi binomiale et tableur

- Calculs et représentations graphiques : utilisation de la fonction LOI.BINOMIALE des tableurs.
- Influence des paramètres n et p sur : l'allure (symétrie ?) et la position du diagramme, la monotonie, la dispersion de X.

Voir grandes binomiales avec tableur | Voir petites binomiales avec tableur

- Allure symétrique "en cloche" centrée sur np des grandes binomiales.
- Déplacement vers la droite du diagramme à n fixé en fonction de la croissance de p.
- Dispersion maximale en $p = 0,5$: étudier la fonction $p \in [0,1] \mapsto np(1-p)$.

Loi binomiale et tableur

- Calculs et représentations graphiques : utilisation de la fonction LOI.BINOMIALE des tableurs.
- Influence des paramètres n et p sur : l'allure (symétrie ?) et la position du diagramme, la monotonie, la dispersion de X.

Voir grandes binomiales avec tableur | Voir petites binomiales avec tableur

- ♣ Allure symétrique "en cloche" centrée sur np des grandes binomiales.
- ♣ Déplacement vers la droite du diagramme à n fixé en fonction de la croissance de p.
- ♣ Dispersion maximale en $p = 0,5$: étudier la fonction $p \in [0,1] \mapsto np(1-p)$.

Loi binomiale et tableur

- Calculs et représentations graphiques : utilisation de la fonction LOI.BINOMIALE des tableurs.
- Influence des paramètres n et p sur : l'allure (symétrie ?) et la position du diagramme, la monotonie, la dispersion de X.

Voir grandes binomiales avec tableur | Voir petites binomiales avec tableur

- ♣ Allure symétrique "en cloche" centrée sur np des grandes binomiales.
- ♣ Déplacement vers la droite du diagramme à n fixé en fonction de la croissance de p.
- ♣ Dispersion maximale en $p = 0,5$: étudier la fonction $p \in [0,1] \mapsto np(1-p)$.

Partie II

La prise de décision

Généralités sur la prise de décision

Logique des règles de décision

1. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ avec certitude.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
2. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ dans 95% des cas.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on ne peut pas conclure sur H_0.
   - Règle de décision **statistique** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
3. Mais en décision statistique, on a 5% de chances de ne pas décider H_0, alors que H_0 est vraie.

Logique des règles de décision

1. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ avec certitude.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
2. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ dans 95% des cas.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on ne peut pas conclure sur H_0.
   - Règle de décision **statistique** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
3. Mais en décision statistique, on a 5% de chances de ne pas décider H_0, alors que H_0 est vraie.

Logique des règles de décision

1. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ avec certitude.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
2. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ dans 95% des cas.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on ne peut pas conclure sur H_0.
   - Règle de décision **statistique** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
3. Mais en décision statistique, on a 5% de chances de ne pas décider H_0, alors que H_0 est vraie.

Logique des règles de décision

1. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ avec certitude.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
2. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ dans 95% des cas.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on ne peut pas conclure sur H_0.
   - Règle de décision **statistique** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
3. Mais en décision statistique, on a 5% de chances de ne pas décider H_0, alors que H_0 est vraie.

Logique des règles de décision

1. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ avec certitude.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
2. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ dans 95% des cas.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on ne peut pas conclure sur H_0.
   - Règle de décision **statistique** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
3. Mais en décision statistique, on a 5% de chances de ne pas décider H_0, alors que H_0 est vraie.

Logique des règles de décision

1. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ avec certitude.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
2. Supposons que le fait qu'une hypothèse H_0 soit vraie implique d'observer un événement $\mathcal{A}$ dans 95% des cas.
   - Règle de décision **déterministe** : Quand on n'observe pas l'événement $\mathcal{A}$, on ne peut pas conclure sur H_0.
   - Règle de décision **statistique** : Quand on n'observe pas l'événement $\mathcal{A}$, on décide de rejeter l'hypothèse H_0.
3. Mais en décision statistique, on a 5% de chances de ne pas décider H_0, alors que H_0 est vraie.

Typologie des erreurs de décision

Dans la prise de décision en faveur de l'une des deux hypothèses H_0 et H_1 (au lycée $H_1 = non\ H_0$), **quatre** cas de figure peuvent se présenter :

	On décide H_0	**On décide H_1**
H_0 **vraie**	Bonne décision	Erreur I
H_0 **fausse**	Erreur II	Bonne décision

Un exemple

- Par exemple, dans un procès, on demande au jury de prendre une décision entre les deux hypothèses :

$$\begin{cases} H_0 & : & \textit{L'accusé est innocent.} \\ H_1 & : & \textit{L'accusé est coupable.} \end{cases}$$

- L'erreur I, dite **de première espèce**, est commise lorsqu'un innocent est déclaré coupable.
- L'erreur II, dite **de seconde espèce**, est commise lorsqu'un coupable est déclaré innocent.

♣ En général, on cherchera à éviter de commettre l'erreur de première espèce (à moins d'avoir une autre philosophie du droit).

Un exemple

- Par exemple, dans un procès, on demande au jury de prendre une décision entre les deux hypothèses :

$$\begin{cases} H_0 & : \quad \textit{L'accusé est innocent.} \\ H_1 & : \quad \textit{L'accusé est coupable.} \end{cases}$$

- L'erreur I, dite **de première espèce**, est commise lorsqu'un innocent est déclaré coupable.
- L'erreur II, dite **de seconde espèce**, est commise lorsqu'un coupable est déclaré innocent.
- ♣ En général, on cherchera à éviter de commettre l'erreur de première espèce (à moins d'avoir une autre philosophie du droit).

Un exemple

- Par exemple, dans un procès, on demande au jury de prendre une décision entre les deux hypothèses :

$$\begin{cases} H_0 & : & \textit{L'accusé est innocent.} \\ H_1 & : & \textit{L'accusé est coupable.} \end{cases}$$

- L'erreur I, dite **de première espèce**, est commise lorsqu'un innocent est déclaré coupable.
- L'erreur II, dite **de seconde espèce**, est commise lorsqu'un coupable est déclaré innocent.
- ♣ En général, on cherchera à éviter de commettre l'erreur de première espèce (à moins d'avoir une autre philosophie du droit).

Un exemple

- Par exemple, dans un procès, on demande au jury de prendre une décision entre les deux hypothèses :

$$\begin{cases} H_0 & : \quad \textit{L'accusé est innocent.} \\ H_1 & : \quad \textit{L'accusé est coupable.} \end{cases}$$

- L'erreur I, dite **de première espèce**, est commise lorsqu'un innocent est déclaré coupable.
- L'erreur II, dite **de seconde espèce**, est commise lorsqu'un coupable est déclaré innocent.
- ♣ En général, on cherchera à éviter de commettre l'erreur de première espèce (à moins d'avoir une autre philosophie du droit).

Remarques sur les erreurs de décision

- Le **risque de première espèce** est la probabilité de ne pas décider H_0, alors que H_0 est vraie. Il est usuellement noté α.
- Le **risque de seconde espèce** est la probabilité de décider H_0, alors que H_1 est vraie. Il est usuellement noté β.
- Les deux types d'erreurs n'ont pas un rôle symétrique.
- En général, on veut surtout éviter de commettre l'erreur de première espèce.
- Le risque de première espèce, α, est très souvent une donnée du problème fixée a priori.
- Au lycée :
  - On fait référence à l'erreur de première espèce avec $\alpha = 5\%$.
  - L'erreur de seconde espèce n'est pas un attendu des programmes.

Remarques sur les erreurs de décision

- Le **risque de première espèce** est la probabilité de ne pas décider H_0, alors que H_0 est vraie. Il est usuellement noté α.
- Le **risque de seconde espèce** est la probabilité de décider H_0, alors que H_1 est vraie. Il est usuellement noté β.
- Les deux types d'erreurs n'ont pas un rôle symétrique.
- En général, on veut surtout éviter de commettre l'erreur de première espèce.
- Le risque de première espèce, α, est très souvent une donnée du problème fixée a priori.
- Au lycée :
  - On fait référence à l'erreur de première espèce avec $\alpha = 5\%$.
  - L'erreur de seconde espèce n'est pas un attendu des programmes.

Remarques sur les erreurs de décision

- Le **risque de première espèce** est la probabilité de ne pas décider H_0, alors que H_0 est vraie. Il est usuellement noté α.
- Le **risque de seconde espèce** est la probabilité de décider H_0, alors que H_1 est vraie. Il est usuellement noté β.
- Les deux types d'erreurs n'ont pas un rôle symétrique.
- En général, on veut surtout éviter de commettre l'erreur de première espèce.
- Le risque de première espèce, α, est très souvent une donnée du problème fixée a priori.
- Au lycée :
  - On fait référence à l'erreur de première espèce avec $\alpha = 5\%$.
  - L'erreur de seconde espèce n'est pas un attendu des programmes.

Remarques sur les erreurs de décision

- Le **risque de première espèce** est la probabilité de ne pas décider H_0, alors que H_0 est vraie. Il est usuellement noté α.
- Le **risque de seconde espèce** est la probabilité de décider H_0, alors que H_1 est vraie. Il est usuellement noté β.
- Les deux types d'erreurs n'ont pas un rôle symétrique.
- En général, on veut surtout éviter de commettre l'erreur de première espèce.
- Le risque de première espèce, α, est très souvent une donnée du problème fixée a priori.
- Au lycée :
  - On fait référence à l'erreur de première espèce avec $\alpha = 5\%$.
  - L'erreur de seconde espèce n'est pas un attendu des programmes.

Remarques sur les erreurs de décision

- Le **risque de première espèce** est la probabilité de ne pas décider H_0, alors que H_0 est vraie. Il est usuellement noté α.
- Le **risque de seconde espèce** est la probabilité de décider H_0, alors que H_1 est vraie. Il est usuellement noté β.
- Les deux types d'erreurs n'ont pas un rôle symétrique.
- En général, on veut surtout éviter de commettre l'erreur de première espèce.
- Le risque de première espèce, α, est très souvent une donnée du problème fixée a priori.
- Au lycée :
  - On fait référence à l'erreur de première espèce avec $\alpha = 5\%$.
  - L'erreur de seconde espèce n'est pas un attendu des programmes.

Remarques sur les erreurs de décision

- Le **risque de première espèce** est la probabilité de ne pas décider H_0, alors que H_0 est vraie. Il est usuellement noté α.
- Le **risque de seconde espèce** est la probabilité de décider H_0, alors que H_1 est vraie. Il est usuellement noté β.
- Les deux types d'erreurs n'ont pas un rôle symétrique.
- En général, on veut surtout éviter de commettre l'erreur de première espèce.
- Le risque de première espèce, α, est très souvent une donnée du problème fixée a priori.
- Au lycée :
  - On fait référence à l'erreur de première espèce avec $\alpha = 5\%$.
  - L'erreur de seconde espèce n'est pas un attendu des programmes.

Remarques sur les erreurs de décision

- Le **risque de première espèce** est la probabilité de ne pas décider H_0, alors que H_0 est vraie. Il est usuellement noté α.
- Le **risque de seconde espèce** est la probabilité de décider H_0, alors que H_1 est vraie. Il est usuellement noté β.
- Les deux types d'erreurs n'ont pas un rôle symétrique.
- En général, on veut surtout éviter de commettre l'erreur de première espèce.
- Le risque de première espèce, α, est très souvent une donnée du problème fixée a priori.
- Au lycée :
  - On fait référence à l'erreur de première espèce avec $\alpha = 5\%$.
  - L'erreur de seconde espèce n'est pas un attendu des programmes.

Remarques sur les erreurs de décision

- Le **risque de première espèce** est la probabilité de ne pas décider H_0, alors que H_0 est vraie. Il est usuellement noté α.
- Le **risque de seconde espèce** est la probabilité de décider H_0, alors que H_1 est vraie. Il est usuellement noté β.
- Les deux types d'erreurs n'ont pas un rôle symétrique.
- En général, on veut surtout éviter de commettre l'erreur de première espèce.
- Le risque de première espèce, α, est très souvent une donnée du problème fixée a priori.
- Au lycée :
  - On fait référence à l'erreur de première espèce avec $\alpha = 5\%$.
  - L'erreur de seconde espèce n'est pas un attendu des programmes.

Prise de décision avec les outils de la Seconde

Intervalle de fluctuation introduit en Seconde : IF-Seconde

- Si $n > 25$ et $0,2 < p < 0,8$, la probabilité, que la **fréquence empirique** F de succès dans un échantillon de taille n soit dans $\left[p - \frac{1}{\sqrt{n}}, p + \frac{1}{\sqrt{n}}\right]$, est supérieure à $0,95$ (**seuil**).
- $\mathbb{P}(\mathcal{A}) \geq 0,95$ où $\mathcal{A}$ désigne l'événement « *F appartient à* $\left[p - \frac{1}{\sqrt{n}}, p + \frac{1}{\sqrt{n}}\right]$ ».

♣ L'intervalle $\left[p - \frac{1}{\sqrt{n}}, p + \frac{1}{\sqrt{n}}\right]$ sera appelé **IF-Seconde** dans la suite de l'exposé.

Justifications de l'IF de Seconde

- **Justification par simulation** de l'IF pour $p = 0,4$ à partir d'observations d'un échantillon de taille 100 : Simulation
- **Justification graphique** de l'IF pour $p = 0,4$ à partir d'un histogramme des fréquences obtenues par simulation de 500 observations d'un échantillon de taille 100 : Histogramme
  - On prend **le plus court des intervalles fermés centrés sur** p qui contiennent au minimum 95% des fréquences observées.
  - On prend **le plus court des intervalles fermés centrés sur** p qui laissent au maximum 5% des fréquences observées à l'extérieur.
  - On prend **le plus court des intervalles fermés** qui laissent au maximum 2,5% des fréquences observées de chaque côté extérieur de l'intervalle.

Justifications de l'IF de Seconde

- **Justification par simulation** de l'IF pour $p = 0,4$ à partir d'observations d'un échantillon de taille 100 : Simulation
- **Justification graphique** de l'IF pour $p = 0,4$ à partir d'un histogramme des fréquences obtenues par simulation de 500 observations d'un échantillon de taille 100 : Histogramme
  - On prend **le plus court des intervalles fermés centrés sur** p qui contiennent au minimum 95% des fréquences observées.
  - On prend **le plus court des intervalles fermés centrés sur** p qui laissent au maximum 5% des fréquences observées à l'extérieur.
  - On prend **le plus court des intervalles fermés** qui laissent au maximum $2,5\%$ des fréquences observées de chaque côté extérieur de l'intervalle.

Justifications de l'IF de Seconde

- **Justification par simulation** de l'IF pour $p = 0,4$ à partir d'observations d'un échantillon de taille 100 : Simulation
- **Justification graphique** de l'IF pour $p = 0,4$ à partir d'un histogramme des fréquences obtenues par simulation de 500 observations d'un échantillon de taille 100 : Histogramme
  - On prend **le plus court des intervalles fermés centrés sur** p qui contiennent au minimum 95% des fréquences observées.
  - On prend **le plus court des intervalles fermés centrés sur** p qui laissent au maximum 5% des fréquences observées à l'extérieur.
  - On prend **le plus court des intervalles fermés** qui laissent au maximum $2,5\%$ des fréquences observées de chaque côté extérieur de l'intervalle.

Justifications de l'IF de Seconde

- **Justification par simulation** de l'IF pour $p = 0,4$ à partir d'observations d'un échantillon de taille 100 : Simulation
- **Justification graphique** de l'IF pour $p = 0,4$ à partir d'un histogramme des fréquences obtenues par simulation de 500 observations d'un échantillon de taille 100 : Histogramme
  - On prend **le plus court des intervalles fermés centrés sur** p qui contiennent au minimum 95% des fréquences observées.
  - On prend **le plus court des intervalles fermés centrés sur** p qui laissent au maximum 5% des fréquences observées à l'extérieur.
  - On prend **le plus court des intervalles fermés** qui laissent au maximum $2,5\%$ des fréquences observées de chaque côté extérieur de l'intervalle.

Justifications de l'IF de Seconde

- **Justification par simulation** de l'IF pour $p = 0,4$ à partir d'observations d'un échantillon de taille 100 : Simulation
- **Justification graphique** de l'IF pour $p = 0,4$ à partir d'un histogramme des fréquences obtenues par simulation de 500 observations d'un échantillon de taille 100 : Histogramme
  * On prend **le plus court des intervalles fermés centrés sur** p qui contiennent au minimum 95% des fréquences observées.
  * On prend **le plus court des intervalles fermés centrés sur** p qui laissent au maximum 5% des fréquences observées à l'extérieur.
  * On prend **le plus court des intervalles fermés** qui laissent au maximum $2,5\%$ des fréquences observées de chaque côté extérieur de l'intervalle.

La démarche de prise de décision en Seconde

- On s'intéresse à une expérience de Bernoulli de paramètre inconnu p et on veut décider si $p = p_0$ où p_0 valeur de référence donnée.
- Démarche de la prise de décision :
  * *On observe un échantillon de taille n et on calcule la fréquence observée f,*
  * *Si $f \notin \left[p_0 - \frac{1}{\sqrt{n}}, p_0 + \frac{1}{\sqrt{n}}\right]$, on décide que $p \neq p_0$,*
  * *Si $f \in \left[p_0 - \frac{1}{\sqrt{n}}, p_0 + \frac{1}{\sqrt{n}}\right]$, on décide que $p = p_0$.*

♣ On a (au plus) 5% de chances de se tromper quand on décide que $p \neq p_0$.

Un exemple : La pièce de Buffon (1777)

- Expérience relatée dans *Essai d'Arithmétique morale* (1777) : Buffon lance 4040 fois une pièce et il obtient 2049 fois "Pile". La pièce est-elle équilibrée ?
- La v.a. $X =$"*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*" (i.e. $p = p_0 = 0,5$).
- Les conditions de validité de l'utilisation de l'IF-Seconde sont vérifiées : $n = 4040 > 25$ et $0,2 < p_0 = 0,5 < 0,8$.
- Diagramme en bâtons de la loi de la v.a. X sous l'hypothèse $p = p_0 = 0,5$: Voir le diagramme en bâtons
- Visualisation de la règle de décision avec l'IF-Seconde et réflexion sur la notion d'IF : Voir la règle

Un exemple : La pièce de Buffon (1777)

- Expérience relatée dans *Essai d'Arithmétique morale* (1777) : Buffon lance 4040 fois une pièce et il obtient 2049 fois "Pile". La pièce est-elle équilibrée ?
- La v.a. $X =$"*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*" (i.e. $p = p_0 = 0,5$).
- Les conditions de validité de l'utilisation de l'IF-Seconde sont vérifiées : $n = 4040 > 25$ et $0,2 < p_0 = 0,5 < 0,8$.
- Diagramme en bâtons de la loi de la v.a. X sous l'hypothèse $p = p_0 = 0,5$: Voir le diagramme en bâtons
- Visualisation de la règle de décision avec l'IF-Seconde et réflexion sur la notion d'IF : Voir la règle

Un exemple : La pièce de Buffon (1777)

- Expérience relatée dans *Essai d'Arithmétique morale* (1777) : Buffon lance 4040 fois une pièce et il obtient 2049 fois "Pile". La pièce est-elle équilibrée ?
- La v.a. X ="*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*" (i.e. $p = p_0 = 0,5$).
- Les conditions de validité de l'utilisation de l'IF-Seconde sont vérifiées : $n = 4040 > 25$ et $0,2 < p_0 = 0,5 < 0,8$.
- Diagramme en bâtons de la loi de la v.a. X sous l'hypothèse $p = p_0 = 0,5$: Voir le diagramme en bâtons
- Visualisation de la règle de décision avec l'IF-Seconde et réflexion sur la notion d'IF : Voir la règle

Un exemple : La pièce de Buffon (1777)

- Expérience relatée dans *Essai d'Arithmétique morale* (1777) : Buffon lance 4040 fois une pièce et il obtient 2049 fois "Pile". La pièce est-elle équilibrée ?
- La v.a. X ="*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*" (i.e. $p = p_0 = 0,5$).
- Les conditions de validité de l'utilisation de l'IF-Seconde sont vérifiées : $n = 4040 > 25$ et $0,2 < p_0 = 0,5 < 0,8$.
- Diagramme en bâtons de la loi de la v.a. X sous l'hypothèse $p = p_0 = 0,5$: Voir le diagramme en bâtons
- Visualisation de la règle de décision avec l'IF-Seconde et réflexion sur la notion d'IF : Voir la règle

Un exemple : La pièce de Buffon (1777)

- Expérience relatée dans *Essai d'Arithmétique morale* (1777) : Buffon lance 4040 fois une pièce et il obtient 2049 fois "Pile". La pièce est-elle équilibrée ?
- La v.a. X ="*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*" (i.e. $p = p_0 = 0,5$).
- Les conditions de validité de l'utilisation de l'IF-Seconde sont vérifiées : $n = 4040 > 25$ et $0,2 < p_0 = 0,5 < 0,8$.
- Diagramme en bâtons de la loi de la v.a. X sous l'hypothèse $p = p_0 = 0,5$: Voir le diagramme en bâtons
- Visualisation de la règle de décision avec l'IF-Seconde et réflexion sur la notion d'IF : Voir la règle

Un exemple : La pièce de Buffon (1777)

- Expérience relatée dans *Essai d'Arithmétique morale* (1777) : Buffon lance 4040 fois une pièce et il obtient 2049 fois "Pile". La pièce est-elle équilibrée ?
- La v.a. $X =$"*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*" (i.e. $p = p_0 = 0,5$).
- Les conditions de validité de l'utilisation de l'IF-Seconde sont vérifiées : $n = 4040 > 25$ et $0,2 < p_0 = 0,5 < 0,8$.
- Diagramme en bâtons de la loi de la v.a. X sous l'hypothèse $p = p_0 = 0,5$: Voir le diagramme en bâtons
- Visualisation de la règle de décision avec l'IF-Seconde et réflexion sur la notion d'IF : Voir la règle

Commentaires

- Dans cette activité, les enjeux de la situation sont les mêmes "à droite" et "à gauche".
- La nature de la situation justifie de partager le risque (5%) de façon égale et bilatérale .
- Ici on souhaite décider entre les deux hypothèses : "*La pièce est équilibrée*" (i.e. $p = p_0 = 0,5$) et "*La pièce n'est pas équilibrée*" (i.e. $p \neq 0,5$).
- Les diagrammes de F et de X sont les mêmes : raisonnements avec X ou F équivalents, mais X plus commode.
- Le diagramme est symétrique et centré sur $np = 2020$.
- Le fait que l'IF soit centré ne joue pas de rôle dans la démarche de prise de décision.

IF-Seconde : aspects mathématiques

- La formulation de l'IF-Seconde est admise ; éventuellement justifiée par simulation.
- La formule de l'IF-Seconde n'est valable que pour les grandes binomiales : $n > 25$ et $0,2 < p < 0,8$.
- Le seuil 95% de l'IF est figé car lié à l'expression de l'IF admise. Il ne peut pas être changé.
- Le risque de 5% est partagé à hauteur de 2,5% de chaque côté de l'IF : deux zones de rejet (risque bilatéral symétrisé).
- Conséquence : L'IF-Seconde est centré sur p_0 car les grandes binomiales sont symétriques et centrées sur leur espérance.
- C'est parce que le risque est symétrisé, que l'IF est centré (non l'inverse).

IF-Seconde : interprétation

- L'IF-Seconde est l'endroit où on s'attend à observer la v.a. de décision F si l'hypothèse $p = p_0$ est bonne.
- La v.a. de décision peut varier dans l'IF : c'est la **fluctuation** due au hasard.
- Si l'hypothèse $p = p_0$ est bonne, la v.a. de décision peut quand même sortir de l'IF, mais ceci est très peu probable (probabilité $\leq 0,05$).
- Si la v.a. de décision sort de l'IF, ce n'est pas uniquement dû au hasard : c'est **significatif** que l'hypothèse $p = p_0$ n'est pas bonne.
- La probabilité de rejeter à tort l'hypothèse $p = p_0$ est $0,05$.

Caractérisation de l'IF-Seconde

Sous les conditions $n > 25$ et $0,2 < p < 0,8$, l'IF-Seconde (au seuil 95%) est *approximativement* **le plus petit intervalle fermé** $[f_a\,, f_b]$, vérifiant les deux relations :

- $\mathbb{P}(F \in [0\,, f_a[) \leq 0,025$
  où $[0\,, f_a[$ est la partie extérieure gauche de l'IF ;
- $\mathbb{P}(F \in]f_b\,, 1]) \leq 0,025$
  où $]f_b\,, 1]$ est la partie extérieure droite de l'IF.

Prise de décision avec la loi binomiale

Complément sur l'IF de Seconde

Les outils de Première permettent de calculer la probabilité $\mathbb{P}\left(F \in \left[p - \frac{1}{\sqrt{n}}\,,\, p + \frac{1}{\sqrt{n}}\right]\right)$.

- Pour $n = 30$ et $p = 0,55$, l'intervalle de fluctuation de Seconde est $[0,3674; 0,7326]$ au seuil 95%.
- Un calcul direct avec la loi binomiale conduit à $\mathbb{P}(0,3674 \leq F \leq 0,7326) = \mathbb{P}(11,02 \leq X \leq 21,98) = 0,935$.
- Bien que dans les conditions de grandes binomiales, ici $\mathbb{P}\left(F \in \left[0.55 - \frac{1}{\sqrt{30}}\,,\, 0.55 + \frac{1}{\sqrt{30}}\right]\right) < 0,95$. Pourquoi ?
- Le programme de Terminale apportera un élément de réponse.

Complément sur l'IF de Seconde

Les outils de Première permettent de calculer la probabilité
$\mathbb{P}\left(F \in \left[p - \frac{1}{\sqrt{n}}\, , \, p + \frac{1}{\sqrt{n}}\right]\right)$.

- Pour $n = 30$ et $p = 0,55$, l'intervalle de fluctuation de Seconde est $[0,3674 ; 0,7326]$ au seuil 95%.
- Un calcul direct avec la loi binomiale conduit à $\mathbb{P}(0,3674 \leq F \leq 0,7326) = \mathbb{P}(11,02 \leq X \leq 21,98) = 0,935$.
- Bien que dans les conditions de grandes binomiales, ici $\mathbb{P}\left(F \in \left[0.55 - \frac{1}{\sqrt{30}}\, , \, 0.55 + \frac{1}{\sqrt{30}}\right]\right) < 0,95$. Pourquoi ?
- Le programme de Terminale apportera un élément de réponse.

Complément sur l'IF de Seconde

Les outils de Première permettent de calculer la probabilité $\mathbb{P}\left(F \in \left[p - \frac{1}{\sqrt{n}}\, , \, p + \frac{1}{\sqrt{n}}\right]\right)$.

- Pour $n = 30$ et $p = 0,55$, l'intervalle de fluctuation de Seconde est $[0,3674; 0,7326]$ au seuil 95%.
- Un calcul direct avec la loi binomiale conduit à $\mathbb{P}(0,3674 \leq F \leq 0,7326) = \mathbb{P}(11,02 \leq X \leq 21,98) = 0,935$.
- Bien que dans les conditions de grandes binomiales, ici $\mathbb{P}\left(F \in \left[0.55 - \frac{1}{\sqrt{30}}\, , \, 0.55 + \frac{1}{\sqrt{30}}\right]\right) < 0,95$. Pourquoi ?
- Le programme de Terminale apportera un élément de réponse.

Complément sur l'IF de Seconde

Les outils de Première permettent de calculer la probabilité $\mathbb{P}\left(F \in \left[p - \frac{1}{\sqrt{n}}, p + \frac{1}{\sqrt{n}}\right]\right)$.

- Pour $n = 30$ et $p = 0,55$, l'intervalle de fluctuation de Seconde est $[0,3674; 0,7326]$ au seuil 95%.
- Un calcul direct avec la loi binomiale conduit à $\mathbb{P}(0,3674 \leq F \leq 0,7326) = \mathbb{P}(11,02 \leq X \leq 21,98) = 0,935$.
- Bien que dans les conditions de grandes binomiales, ici $\mathbb{P}\left(F \in \left[0.55 - \frac{1}{\sqrt{30}}, 0.55 + \frac{1}{\sqrt{30}}\right]\right) < 0,95$. Pourquoi ?
- Le programme de Terminale apportera un élément de réponse.

Probabilité de recouvrement de l'IF-Seconde

- La probabilité $\mathcal{C}_2(n,p) = \mathbb{P}\left(F \in \left[p - \frac{1}{\sqrt{n}}\,,\, p + \frac{1}{\sqrt{n}}\right]\right)$ est dite **de recouvrement** de l'IF-Seconde.
- Notons que $\mathcal{C}_2(n,p) = \mathbb{P}\left(X \in \left[np - \sqrt{n}, np + \sqrt{n}\right]\right)$, d'où :

$$\mathcal{C}_2(n,p) = \sum_{k \in \mathbb{N} \cap \left[np - \sqrt{n}, np + \sqrt{n}\right]} \binom{n}{k} p^k (1-p)^{n-k} .$$

- Variations de $\mathcal{C}_2(n,p)$ en fonction de n et p : Graphiques .

Probabilité de recouvrement de l'IF-Seconde

- La probabilité $\mathcal{C}_2(n,p) = \mathbb{P}\left(F \in \left[p - \frac{1}{\sqrt{n}}\,,\, p + \frac{1}{\sqrt{n}}\right]\right)$ est dite **de recouvrement** de l'IF-Seconde.
- Notons que $\mathcal{C}_2(n,p) = \mathbb{P}\left(X \in \left[np - \sqrt{n}, np + \sqrt{n}\right]\right)$, d'où :

$$\mathcal{C}_2(n,p) = \sum_{k \in \mathbb{N} \cap \left[np - \sqrt{n}, np + \sqrt{n}\right]} \binom{n}{k} p^k (1-p)^{n-k} .$$

- Variations de $\mathcal{C}_2(n,p)$ en fonction de n et p : Graphiques .

Probabilité de recouvrement de l'IF-Seconde

- La probabilité $\mathcal{C}_2(n,p) = \mathbb{P}\left(F \in \left[p - \frac{1}{\sqrt{n}}\,,\, p + \frac{1}{\sqrt{n}}\right]\right)$ est dite **de recouvrement** de l'IF-Seconde.
- Notons que $\mathcal{C}_2(n,p) = \mathbb{P}\left(X \in \left[np - \sqrt{n}, np + \sqrt{n}\right]\right)$, d'où :

$$\mathcal{C}_2(n,p) = \sum_{k \in \mathbb{N} \cap \left[np - \sqrt{n}, np + \sqrt{n}\right]} \binom{n}{k} p^k (1-p)^{n-k} \,.$$

- Variations de $\mathcal{C}_2(n,p)$ en fonction de n et p : Graphiques .

Buffon : calculs avec la loi binomiale

- La variable aléatoire "*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*". Les calculs sont conduits avec $p = p_0 = 0,5$.
- Raisonnement avec X et prise de décision en exploitant la loi binomiale sous l'hypothèse $p = p_0 = 0,5$:
  Raisonner avec les probabilités cumulées
- Visualisation de la règle de décision : Voir la règle

Buffon : calculs avec la loi binomiale

- La variable aléatoire "*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*". Les calculs sont conduits avec $p = p_0 = 0,5$.
- Raisonnement avec X et prise de décision en exploitant la loi binomiale sous l'hypothèse $p = p_0 = 0,5$:
  Raisonner avec les probabilités cumulées
- Visualisation de la règle de décision : Voir la règle

Buffon : calculs avec la loi binomiale

- La variable aléatoire "*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*". Les calculs sont conduits avec $p = p_0 = 0,5$.
- Raisonnement avec X et prise de décision en exploitant la loi binomiale sous l'hypothèse $p = p_0 = 0,5$:
  Raisonner avec les probabilités cumulées
- Visualisation de la règle de décision : Voir la règle

Buffon : calculs avec la loi binomiale

- La variable aléatoire "*Nombre de 'Pile' obtenu sur les* 4040 *lancers*" suit la loi binomiale $\mathcal{B}(n = 4040\,;\, p)$, où p est inconnu.
- L'hypothèse à tester est "*La pièce est équilibrée*". Les calculs sont conduits avec $p = p_0 = 0,5$.
- Raisonnement avec X et prise de décision en exploitant la loi binomiale sous l'hypothèse $p = p_0 = 0,5$:
  Raisonner avec les probabilités cumulées
- Visualisation de la règle de décision : Voir la règle

Comparaison avec la Seconde

- L'IF bilatéral pour X, $[1958\,;\,2082]$ (resp. pour F, $[0,4847\,;\,0,5153]$) est obtenu par un calcul exact (IF-exact). Il est centré sur la valeur $np = 2020$ (resp. $p = 0,5$) ;
- mais le fait d'être centré n'est pas intervenu dans le raisonnement.
- L'IF exact $[0,4847\,;\,0,5153]$ est strictement inclus dans l'IF-Seconde $[0,4843\,;\,0,5157]$.
- Ici, l'IF bilatéral calculé avec la loi binomiale pour F est très proche de l'IF-Seconde.
- Les deux méthodes de prise de décision (avec IF-exact et IF-Seconde) conduisent ici pratiquement à la même décision.
- ♣ Attention ! Si observation près des bornes, c'est l'IF-exact (et non l'IF-Seconde) qu'il faut utiliser pour décider.

Synthèse sur l'IF exact de Première

IF bilatéral exact : définition élémentaire

La variable de décision considérée ici est l'effectif empirique $X = nF$ de loi $\mathcal{B}(n\,;\,p_0)$ où p_0 est donnée.

On appellera **intervalle de fluctuation bilatéral exact (de Première) pour X au seuil** 95% le plus petit intervalle fermé $[a_n\,,\,b_n]$ qui laisse au maximum $2,5\%$ des fréquences observées de chaque côté extérieur de l'intervalle.

IF bilatéral exact : définition équivalente

La variable de décision considérée ici est l'effectif empirique $X = nF$ de loi $\mathcal{B}(n\,;\,p_0)$ où p_0 est donnée.

- On montre que l'intervalle de fluctuation bilatéral exact (de Première) pour X au seuil 95% est l'intervalle $[a_n\,,\,b_n]$, où
  - a_n est le plus petit entier tel que $\mathbb{P}(X \leq a_n) > 0,025$;
  - b_n est le plus petit entier tel que $\mathbb{P}(X \leq b_n) \geq 0,975$.
- Faire attention aux inégalités $>$ et $\geq$!
- L'IF bilatéral exact pour la fréquence empirique F au seuil 95% est alors $\left[\dfrac{a_n}{n}\,,\,\dfrac{b_n}{n}\right]$.
- Les entiers a_n et b_n dépendent de n, mais aussi de p.
- Autre nom : IF de **Clopper-Pearson** car il a été étudié par C.-J. Clopper et E.-S. Pearson (1934).

Probabilité de recouvrement de l'IF exact de Première

- La probabilité $\mathcal{C}_E(n,p) = \mathbb{P}\left(F \in \left[\frac{a_n}{n}, \frac{b_n}{n}\right]\right)$, où $\left[\frac{a_n}{n}, \frac{b_n}{n}\right]$ est l'IF exact de Première, est dite **probabilité de recouvrement** de l'IF.
- Notons que $\mathcal{C}_E(n,p) = \mathbb{P}(X \in [a_n, b_n])$, d'où :

$$\mathcal{C}_E(n,p) = \sum_{k \in \mathbb{N} \cap [a_n, b_n]} \binom{n}{k} p^k (1-p)^{n-k} .$$

- Variations de $\mathcal{C}_E(n,p)$ en fonction de n et p : Graphiques .

Probabilité de recouvrement de l'IF exact de Première

- La probabilité $\mathcal{C}_E(n,p) = \mathbb{P}\left(F \in \left[\frac{a_n}{n}, \frac{b_n}{n}\right]\right)$, où $\left[\frac{a_n}{n}, \frac{b_n}{n}\right]$ est l'IF exact de Première, est dite **probabilité de recouvrement** de l'IF.
- Notons que $\mathcal{C}_E(n,p) = \mathbb{P}(X \in [a_n, b_n])$, d'où :

$$\mathcal{C}_E(n,p) = \sum_{k \in \mathbb{N} \cap [a_n, b_n]} \binom{n}{k} p^k (1-p)^{n-k} .$$

- Variations de $\mathcal{C}_E(n,p)$ en fonction de n et p : Graphiques .

Probabilité de recouvrement de l'IF exact de Première

- La probabilité $\mathcal{C}_E(n,p) = \mathbb{P}\left(F \in \left[\frac{a_n}{n}, \frac{b_n}{n}\right]\right)$, où $\left[\frac{a_n}{n}, \frac{b_n}{n}\right]$ est l'IF exact de Première, est dite **probabilité de recouvrement** de l'IF.
- Notons que $\mathcal{C}_E(n,p) = \mathbb{P}(X \in [a_n, b_n])$, d'où :

$$\mathcal{C}_E(n,p) = \sum_{k \in \mathbb{N} \cap [a_n, b_n]} \binom{n}{k} p^k (1-p)^{n-k} \ .$$

- Variations de $\mathcal{C}_E(n,p)$ en fonction de n et p : Graphiques .

Comparaison graphique des IF de Seconde et de Première

- Représentation graphique des bornes des IF de Seconde et de Première en fonctions de p (p en abscisse) : notion d'**abaques** (n.m.).
- Abaques pour l'IF de Seconde et l'IF exact de Première : Voir les graphiques
- Abaques des formulaires de statistique : Voir les graphiques

Comparaison graphique des IF de Seconde et de Première

- Représentation graphique des bornes des IF de Seconde et de Première en fonctions de p (p en abscisse) : notion d'**abaques** (n.m.).
- Abaques pour l'IF de Seconde et l'IF exact de Première : Voir les graphiques
- Abaques des formulaires de statistique : Voir les graphiques

Comparaison graphique des IF de Seconde et de Première

- Représentation graphique des bornes des IF de Seconde et de Première en fonctions de p (p en abscisse) : notion d'**abaques** (n.m.).
- Abaques pour l'IF de Seconde et l'IF exact de Première : Voir les graphiques
- Abaques des formulaires de statistique : Voir les graphiques

IF exact et prise de décision en Première

La v.a. de décision considérée est $X = nF$ de loi $\mathcal{B}(n\,;\,p_0)$ où p_0 est donnée.
Suivant la même démarche qu'en Seconde, si $[a_n\,,\,b_n]$ désigne l'intervalle de fluctuation bilatéral exact pour l'effectif empirique $X := nF$ au seuil 95%, on lui associera la méthode suivante de prise de décision :
Pour décider si $p = p_0$, on prélève un échantillon de taille n et on calcule la fréquence observée $f = \dfrac{x}{n}$, où x est l'effectif (nombre de succès) observé :

- *Si* $f \notin \left[\dfrac{a_n}{n}\,,\,\dfrac{b_n}{n}\right]$ *(i.e.* $x \notin [a_n\,,\,b_n]$*), on décide que* $p \neq p_0$.
- *Si* $f \in \left[\dfrac{a_n}{n}\,,\,\dfrac{b_n}{n}\right]$ *(i.e.* $x \in [a_n\,,\,b_n]$*), on décide que* $p = p_0$.

Commentaires sur l'IF exact

- Comme $\mathbb{P}\left(F \in \left[\frac{a_n}{n}, \frac{b_n}{n}\right]\right) = \mathbb{P}\left(X \in [a_n\,,\, b_n]\right)$ où $X = nF$ est une v.a. binomiale $\mathcal{B}(n, p_0)$, on peut raisonner avec X ou F et les IF correspondants.
- Cette méthode,
  * est conduite avec des calculs exacts ;
  * est valable pour toute binomiale avec n et p quelconques ;
  * se généralise aisément au cas où le seuil choisi n'est plus 95% ;
  * englobe la méthode vue en Seconde.

Partie III

Bilatéral ou unilatéral ?

L'affaire Woburn

L'affaire Woburn : cadre probabiliste

- L'expérience aléatoire de départ, $\mathcal{E}$: "*Extraire une boule de l'urne et noter sa couleur*", est une expérience de Bernoulli avec $p = 0,00052$.
- L'expérience étudiée est l'expérience $\mathcal{E}_{5969}$ obtenue par 5969 "répétitions" à l'identique de $\mathcal{E}$.
- On associe à l'expérience aléatoire $\mathcal{E}_{5969}$ un modèle probabiliste $(\Omega, \mathbb{P})$ et la variable aléatoire X qui, à toute issue ω de l'expérience $\mathcal{E}_{5969}$ fait correspondre $X(\omega) :=$*nombre de boules rouges comptées dans l'issue* ω.

L'affaire Woburn : cadre probabiliste

- L'expérience aléatoire de départ, $\mathcal{E}$: "*Extraire une boule de l'urne et noter sa couleur*", est une expérience de Bernoulli avec $p = 0,00052$.
- L'expérience étudiée est l'expérience $\mathcal{E}_{5969}$ obtenue par 5969 "répétitions" à l'identique de $\mathcal{E}$.
- On associe à l'expérience aléatoire $\mathcal{E}_{5969}$ un modèle probabiliste $(\Omega, \mathbb{P})$ et la variable aléatoire X qui, à toute issue ω de l'expérience $\mathcal{E}_{5969}$ fait correspondre $X(\omega) :=$ *nombre de boules rouges comptées dans l'issue* ω.

L'affaire Woburn : cadre probabiliste

- L'expérience aléatoire de départ, $\mathcal{E}$: "*Extraire une boule de l'urne et noter sa couleur*", est une expérience de Bernoulli avec $p = 0,00052$.
- L'expérience étudiée est l'expérience $\mathcal{E}_{5969}$ obtenue par 5969 "répétitions" à l'identique de $\mathcal{E}$.
- On associe à l'expérience aléatoire $\mathcal{E}_{5969}$ un modèle probabiliste $(\Omega, \mathbb{P})$ et la variable aléatoire X qui, à toute issue ω de l'expérience $\mathcal{E}_{5969}$ fait correspondre $X(\omega) :=$*nombre de boules rouges comptées dans l'issue* ω.

L'affaire Woburn : détermination de la probabilité

- On s'intéresse à la probabilité, $\mathbb{P}(X = 9)$, d'obtenir 9 boules rouges lorsqu'on effectue 5969 épreuves de l'expérience "*Extraire une boule de l'urne et noter sa couleur*".
- L'expression de la probabilité cherchée est $\mathbb{P}(X = 9) = \binom{5969}{9} \times 0,00052^9 \times 0,99948^{5960}$. Valeur ?
- Utilisation d'un tableur : Détermination de $\mathbb{P}(X = 9) \simeq 0,0033$. Voir diagramme Woburn

L'affaire Woburn : détermination de la probabilité

- On s'intéresse à la probabilité, $\mathbb{P}(X=9)$, d'obtenir 9 boules rouges lorsqu'on effectue 5969 épreuves de l'expérience "*Extraire une boule de l'urne et noter sa couleur*".
- L'expression de la probabilité cherchée est $\mathbb{P}(X=9)=\binom{5969}{9}\times 0,00052^9\times 0,99948^{5960}$. Valeur ?
- Utilisation d'un tableur : Détermination de $\mathbb{P}(X=9)\simeq 0,0033$. Voir diagramme Woburn

L'affaire Woburn : détermination de la probabilité

- On s'intéresse à la probabilité, $\mathbb{P}(X = 9)$, d'obtenir 9 boules rouges lorsqu'on effectue 5969 épreuves de l'expérience "*Extraire une boule de l'urne et noter sa couleur*".
- L'expression de la probabilité cherchée est $\mathbb{P}(X = 9) = \binom{5969}{9} \times 0,00052^9 \times 0,99948^{5960}$. Valeur ?
- Utilisation d'un tableur : Détermination de $\mathbb{P}(X = 9) \simeq 0,0033$. Voir diagramme Woburn

Réflexions sur l'affaire Woburn

- La variable aléatoire "*Nombre de cas de leucémie obtenu sur les* 5969 *adolescents*" suit la loi binomiale $\mathcal{B}(n = 5969\,;\ p)$, où p est inconnu.
- Deux problématiques pour le décideur :
  1. La situation est-elle normale à Woburn ?
  2. La situation est-elle dangereuse pour la santé à Woburn ?

Réflexions sur l'affaire Woburn

- La variable aléatoire "*Nombre de cas de leucémie obtenu sur les* 5969 *adolescents*" suit la loi binomiale $\mathcal{B}(n = 5969\,;\ p)$, où p est inconnu.
- Deux problématiques pour le décideur :
  1. La situation est-elle normale à Woburn ?
  2. La situation est-elle dangereuse pour la santé à Woburn ?

Réflexions sur l'affaire Woburn

- La variable aléatoire "*Nombre de cas de leucémie obtenu sur les* 5969 *adolescents*" suit la loi binomiale $\mathcal{B}(n = 5969\,;\, p)$, où p est inconnu.
- Deux problématiques pour le décideur :
  1. La situation est-elle normale à Woburn ?
  2. La situation est-elle dangereuse pour la santé à Woburn ?

Réflexions sur l'affaire Woburn

- La variable aléatoire "*Nombre de cas de leucémie obtenu sur les* 5969 *adolescents*" suit la loi binomiale $\mathcal{B}(n = 5969\,;\ p)$, où p est inconnu.
- Deux problématiques pour le décideur :
  1. La situation est-elle normale à Woburn ?
  2. La situation est-elle dangereuse pour la santé à Woburn ?

La situation est-elle normale à Woburn ?

- Raisonnement avec X et prise de décision (bilatérale) en exploitant la loi binomiale avec $p = p_0 = 0,00052$:

Raisonner avec les probabilités cumulées

- Visualiser les règles de décision bilatérale (à 95% et 90%) :

Visualisation

La situation est-elle normale à Woburn ?

- Raisonnement avec X et prise de décision (bilatérale) en exploitant la loi binomiale avec $p = p_0 = 0,00052$:

Raisonner avec les probabilités cumulées

- Visualiser les règles de décision bilatérale (à 95% et 90%) :

Visualisation

La situation est-elle dangereuse à Woburn ?

- Raisonnement avec X et prise de décision (unilatérale) en exploitant la loi binomiale avec $p = p_0 = 0,00052$:

  Raisonner avec les probabilités cumulées

- Visualiser la règle de décision unilatérale :

  Visualisation

La situation est-elle dangereuse à Woburn ?

- Raisonnement avec X et prise de décision (unilatérale) en exploitant la loi binomiale avec $p = p_0 = 0,00052$:

Raisonner avec les probabilités cumulées

- Visualiser la règle de décision unilatérale :

Visualisation

Commentaires

- Le choix des hypothèses est effectué en premier. Il dépend des préoccupations du décideur liées aux enjeux (économiques, sociaux, sanitaires, politiques, ...) de la situation.
- Le choix des hypothèses induit ensuite la forme de la zone de rejet, et donc la forme de l'IF.
- Il est nécessaire de faire une analyse de la signification concrète des risques encourus.
- Dans cette activité, les erreurs ne correspondent pas au même enjeu "à droite" et "à gauche".
- La nature de la situation conduit donc plutôt à envisager un risque unilatéral.

Choix des hypothèses

- Soit on souhaite décider entre "*La situation est normale*" (i.e. $p = 0.00052$) et "*La situation n'est pas normale*" (i.e. $p \neq 0.00052$) : risque bilatéral.
- Soit on souhaite décider entre "*La situation présente un danger pour la santé*" (i.e. $p > 0,00052$) et "*La situation ne présente pas un danger pour la santé*" (i.e. $p \leq 0,00052$) : risque unilatéral.
- Dans les deux cas les calculs sont effectués avec $p = 0,00052$.
- La variable de décision utilisée est $X = nF$ de loi binomiale $\mathcal{B}(n = 5969\,;\ p = 0,00052)$.

Décisions et choix du seuil

- Comme l'observation donne $x = 9$, on décide que la situation est, avec le même seuil 95%,
  - anormale en utilisant l'IF bilatéral,
  - dangereuse en utilisant l'IF unilatéral.
- Il y a 5% de chances de se tromper en prenant chacune de ces décisions.
- Si l'observation avait donné $x = 0$ ou $x = 7$, on aurait été amené à déclarer la situation :
  - anormale avec l'IF bilatéral au seuil 90%,
  - mais normale avec l'IF bilatéral au seuil 95%.
- Avec le même seuil 95%, si l'observation avait donné $x = 7$, on aurait été amené à déclarer la situation :
  - normale en utilisant l'IF bilatéral,
  - mais dangereuse en utilisant l'IF unilatéral.

Remarques sur l'IF exact et l'IF-Seconde

- Ici l'IF exact bilatéral $[0\,;\,0,0012]$ déterminés au seuil 95% n'est pas centré sur $p_0 = 0,00052$.
- L'IF exact bilatéral $[0\,;\,0,0012]$ déterminés au seuil 95% est nettement différent de l'IF-Seconde :

$$\left[0,00052 - \frac{1}{\sqrt{5969}}\,;\,0,00052 + \frac{1}{\sqrt{5969}}\right] = [-0,0124\,;\,0,0135]\ .$$

- Il serait aberrant d'utiliser l'IF-Seconde dans le contexte de cette situation pour prendre la décision (car petite binomiale).

Conclusion

Sur la démarche de prise de décision en général

Trois étapes :

1. **L'analyse concrète de la situation** qui conduit à préciser l'hypothèse à tester et les contraintes de l'intervalle de fluctuation.
2. **La formalisation théorique** qui conduit à la détermination de l'intervalle de fluctuation et à l'énoncé de la règle de décision.
3. **L'observation d'un échantillon et la décision** prise par l'application de la règle de décision à cette observation.

Sur la démarche de prise de décision en général

Trois étapes :

1. **L'analyse concrète de la situation** qui conduit à préciser l'hypothèse à tester et les contraintes de l'intervalle de fluctuation.
2. **La formalisation théorique** qui conduit à la détermination de l'intervalle de fluctuation et à l'énoncé de la règle de décision.
3. **L'observation d'un échantillon et la décision** prise par l'application de la règle de décision à cette observation.

Sur la démarche de prise de décision en général

Trois étapes :

1. **L'analyse concrète de la situation** qui conduit à préciser l'hypothèse à tester et les contraintes de l'intervalle de fluctuation.
2. **La formalisation théorique** qui conduit à la détermination de l'intervalle de fluctuation et à l'énoncé de la règle de décision.
3. **L'observation d'un échantillon et la décision** prise par l'application de la règle de décision à cette observation.

Sur les raisonnements utilisés

- Réinvestissement de la Troisième et de la Seconde (expérience aléatoire, arbres, modèles, événement, prise de décision, ...).
- Justification de la règle de multiplication de probabilités sur les arbres.
- Travail sur la notion de v.a. et de loi de probabilité.
- Possibilité de mener des calculs exacts avec la loi binomiale.
- Importance de la signification concrète de la situation, et du regard du décideur sur celle-ci, dans la formulation des hypothèses, l'analyse des enjeux et la forme de l'IF.

Bibliographie

Bibliographie

Ducel Y., Saussereau B.,
Quelle problématique pour un enseignement des probabilités en Troisième ?
Repères IREM, 77, octobre 2009, pages 55-65, Topiques éditions, Nancy, 2009.

Ducel Y.,
Les probabilités en Troisième et en Seconde : mise en perspective des programmes
Diaporama, Atelier, 24 octobre 2010, Journées nationales de l'APMEP, Paris, 2010.

Bibliographie

Ducel Y., Saussereau B.,
Quelle problématique pour un enseignement des probabilités en Troisième ?
Repères IREM, 77, octobre 2009, pages 55-65, Topiques éditions, Nancy, 2009.

Ducel Y.,
Les probabilités en Troisième et en Seconde : mise en perspective des programmes
Diaporama, Atelier, 24 octobre 2010, Journées nationales de l'APMEP, Paris, 2010.

Bibliographie

- Ducel Y., Saussereau B.,
  La prise de décision de la Seconde à la Première
  Repères IREM, 85, octobre 2011, pages 31-49, Topiques éditions, Nancy, 2011 (consultable en ligne sur le portail des IREM).

- Ministère Éducation nationale,
  Statistiques et probabilités : classe de Première générale et technologique
  Document ressource, Ministère de l'Éducation nationale, mai 2011 (site Web du groupe des mathématiques de l'inspection générale, http ://igmaths.infos.st/spip/).

Bibliographie

Ducel Y., Saussereau B.,
La prise de décision de la Seconde à la Première
Repères IREM, 85, octobre 2011, pages 31-49, Topiques éditions, Nancy, 2011 (consultable en ligne sur le portail des IREM).

Ministère Éducation nationale,
Statistiques et probabilités : classe de Première générale et technologique
Document ressource, Ministère de l'Éducation nationale, mai 2011 (site Web du groupe des mathématiques de l'inspection générale, http ://igmaths.infos.st/spip/).

Merci de votre attention

Pour nous contacter ...

- **Yves Ducel, Bruno Saussereau**
  Institut de recherche sur l'enseignement des mathématiques (IREM)

- **Téléphones** : (YD) +33(0)3 81 66 62 32 / (BS) +33(0)3 81 66 63 00
  Adresses électroniques : yves.ducel@univ-fcomte.fr / bruno.saussereau@univ-fcomte.fr

- **Adresse postale** : IREM - Département de mathématiques
  UFR Sciences et techniques de l'Université de Franche-Comté
  16, route de Gray, F-25030 Besançon cedex

♣ **Site Web de l'IREM** : http ://www-irem.univ-fcomte.fr/