

## FORMULES

- Objectif : estimer  $\pi = \sum_{i=1}^N \frac{Y_i}{N}$
- $\pi \sim \text{Beta}(a, b)$ ; Uniforme si  $a = b = 1$
- Si  $t$  est le nombre de succès observés et a priori  $\pi \sim \text{Beta}(a, b)$  :  
A posteriori :  $(\pi|T = t) \sim \text{Beta}(a + t, b + n - t) \approx N(\hat{\pi}, \frac{\hat{\pi}(1-\hat{\pi})}{n})$  avec  $\hat{\pi} = \frac{(a-1)+t}{(a+b-2)+n}$
- Si  $\delta = \pi_2 - \pi_1, (\delta = \pi_2 - \pi_1 | T_1 = t_1, T_2 = t_2) \approx N(\hat{\pi}_2 - \hat{\pi}_1, \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2} + \frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1})$

## EXERCICES À FAIRE À LA SÉANCE DE TP

- En temps normal dans une usine de production de bois qui possède 250 machines, la probabilité qu'une machine tombe en panne dans une journée de travail ( $\pi$ ) est de 0,01. Un jour, on a observé que parmi 250 appareils, 4 appareils sont tombés en panne.
  - (Avec un a priori uniforme sur  $\pi$ ), déterminez la densité a posteriori de  $\pi$  connaissant cet événement.
  - Déterminez si cet événement indique que la production n'est plus sous contrôle. (Suggestion: utilisez une approximation normale de la distribution a posteriori pour calculer la probabilité a posteriori que  $\pi > 0.01$ ).
  - Déterminez un ensemble de valeurs plausibles pour  $\pi$ .
- L'association des fabricants d'automobiles des Etats-Unis relate dans son périodique 'Motor Vehicle Facts and Figures' que, en 1962, parmi toutes les ventes de voitures aux USA, 4.8% sont des voitures étrangères. En 1983, le pourcentage était 27.8%. Du coup, les fabricants américains ont implémenté différentes règles permettant de réduire les coûts de manière à concurrencer les voitures d'importation. Parmi un échantillon de 500 voitures vendues en 1991, 128 sont des voitures d'importation.
  - Ces données suggèrent-elles que le pourcentage de voitures étrangères vendues a diminué entre 1983 et 1991 où ce pourcentage valait 27.8%? (suggestion: approximer la probabilité adéquate afin d'argumenter votre réponse)
  - Déterminer un ensemble de valeurs plausibles pour la proportion de voitures étrangères vendues en 1991.
  - Supposons que cette proportion de voitures d'importation vendues aux USA soit de nouveau étudiée sur l'année 1992 et que l'on obtienne que parmi un échantillon de 800 voitures, 172 sont des voitures d'importation. Cette nouvelle donnée suggère-t-elle que le pourcentage de voitures étrangères vendues a diminué entre 1983 et 1992 ?
- Aux Etats-Unis le « U.S. Energy Information Administration » effectue des études afin d'estimer le pourcentage de ménages américains possédant différents appareils. Les résultats de ces études peuvent être trouvés dans « Residential Energy Consumption Survey: Housing Characteristics ». Euromonitor Publications Limited, London, England conduit des études similaires dans d'autres pays et publie leurs résultats dans « European Marketing Data and Statistics ». Supposons que

hors d'un échantillon aléatoire de 500 ménages Américains, 370 possèdent leur propre machine à lessiver et que hors d'un échantillon aléatoire de 450 ménages Français, 365 possèdent cet appareil.

- a. Ces données suggèrent-elles une différence entre les pourcentages de ménages Américains et Français possédant une machine à lessiver? (suggestion: déterminer la distribution a posteriori de la différence des pourcentages, et calculer la probabilité a posteriori que cette différence soit positive ou négative)
  - b. Donner un ensemble de valeurs plausibles pour cette différence de pourcentages.
4. Une très grande compagnie américaine emploie 75% d'hommes et 25% de femmes. En 1988, 19.1% des employés ont été promus. Au début de l'année suivante, une des employées a porté plainte contre la compagnie, invoquant une discrimination envers les femmes sur base que seulement 61 des 272 employés promus l'année précédente étaient des femmes.
- a. Quels pourcentages le Juge va-t-il devoir comparer pour trancher le litige ?
  - b. N'ayant pas d'autres informations sur les proportions envisagées dans ce problème, déterminer la conclusion que le juge devrait porter à partir de l'argument invoqué ? (suggestion: déterminer la distribution a posteriori de la probabilité qu'une femme soit promue et calculer, éventuellement à l'aide d'une approximation, une probabilité a posteriori pertinente pour répondre à la question.)
  - c. Donner un ensemble de valeurs plausibles pour cette probabilité.

Si en 1989, parmi 176 employés promus, 17.6% sont des femmes :

- d. Déterminer s'il y a lieu maintenant de considérer que cette compagnie est discriminante envers les femmes ? (suggestion: déterminer une nouvelle distribution approximative tenant compte des informations disponibles sur les 2 années.)
- e. Que devient l'ensemble des valeurs plausibles pour cette probabilité ?

Si en 2007, parmi 146 employés promus, 13.7% sont des femmes:

- f. Déterminer si la politique de promotion de l'entreprise a changé depuis 1989 ? (suggestion: déterminer la distribution a posteriori de la différence des pourcentages de femmes promues et calculer, éventuellement à l'aide d'une approximation, une probabilité a posteriori pertinente pour répondre à la question)
- g. Donner un ensemble de valeurs plausibles pour cette probabilité.

5. Une autre compagnie américaine emploie 150000 personnes dont 85750 hommes. En 2006, 2.3% des employés ont été promus. Au début de l'année suivante, une des employées a porté plainte contre la compagnie, invoquant une discrimination envers les femmes sur base que seulement 1400 des 3450 employés promus l'année précédente étaient des femmes.

- a. Quels pourcentages le Juge va-t-il devoir comparer pour trancher le litige ?
- b. N'ayant pas d'autres informations sur les proportions envisagées dans ce problème, déterminer la conclusion que le juge devrait porter à partir de l'argument invoqué ? (suggestion: déterminer la distribution a posteriori d'une différence de probabilités et calculer une probabilité a posteriori pertinente pour répondre à la question.)
- c. Donner un ensemble de valeurs plausibles pour cette probabilité.

## EXERCICES COMPLÉMENTAIRES

1. Il est communément admis que dans une bibliothèque, environ 10% des livres se détériorent chaque année. Le suivi spécifique de 400 livres durant une année a révélé 36 détériorations.
  - a. Que vaut la probabilité (a posteriori) qu'au plus 10% des livres de la bibliothèque ont subi des détériorations cette année-là ?
  - b. Donnez un ensemble de valeurs plausibles pour cette probabilité.

Supposons que l'année suivante sur un échantillon de 300 livres suivis, on sache que 6% de livres aient subi des détériorations.

- c. En supposant que le pourcentage de détérioration soit stable sur les 2 années, répondez de nouveau aux questions a) et b).
2. Supposons que le taux de mortalité, calculé dans un échantillon d'effectif 100 de personnes touchées par une certaine maladie, soit de 0,13. Soit  $\pi$ , cette probabilité de décès au niveau de la population.
  - a. Ces données suggèrent-elles que la probabilité de mourir pour toutes les personnes touchées par cette maladie ( $\pi$ ) soit supérieure à 0,1?
  - b. Répondre à la question en raisonnant à partir d'un ensemble de valeurs plausibles pour cette probabilité.
3. Un vaccin antigrippal classique est administré à 50 personnes sélectionnées aléatoirement. Dix d'entre elles sont néanmoins grippées l'hiver suivant. Simultanément, un nouveau vaccin est mis sur le marché et testé sur 50 personnes sélectionnées également de façon aléatoire. Cinq parmi celles-ci sont grippées l'hiver suivant. Peut-on affirmer que le nouveau vaccin est plus efficace? Répondez à la question de deux façons différentes.