IUFM de Midi-Pyrénées
Prparation au CAPES de Maths
2000-2001

Feuille d'exercices 6

1 Statistique sur le paramètre d'une binomiale

1.1 Contrôle de qualité

Une entreprise fabrique une pièce de moteurs industriels. Parfois ces pièces se révèlent immédiatement défaillantes après la vente. Le taux de défaillance doit être limité à 4%. Sur 500 pièces contrôlées, 28 sont défaillantes.

- Donner une fourchette de confiance pour le taux en question (risque 5%).
- Est-ce-que la norme de qualité de production est respectée?

1.2 Bonbecs

Un expérimentateur offre quatres types de bonbons de différentes tailles à des personnes en leur demandant d'en prendre juste un. Parmi 100 enfants tentés, 60 ont choisi un bonbon de la plus grande taille. Tandis que parmi 120 adultes 60 l'ont choisi. Est-il évident que la tentation de "grand bonbon" est différente chez les enfants et chez les adultes?

1.3 Lumineux

Des ampoules provenant d'une certaine fabrication peuvent fonctionner plus de 200 heures. A la suite d'un nouveau traitement appliqué au filament, on constate, sur un échantillon de 100 ampoules, que 30 d'entre elles fonctionnent plus de 200 heures. L'amélioration apparente est-elle significative au seuil 5%?

1.4 Traitement

On compare les effets d'un même traitement dans deux hôpitaux différents. Dans le premier hôpital, sur 100 malades traités, 70 montrent des signes de guérison alors que dans le second hôpital, sur 150 malades traités, 100 sont sur le point de guérir. Quelle conclusion peut-on en tirer, au risque 5%?

1.5 La pêche miraculeuse

Dans un étang, il y a deux sortes de poissons, des gardons et des carpes. On appelle p la proportion de gardons dans l'étang. On suppose que $p = 0,3$.

- Un pêcheur attrape 8 poissons. On note X le nombre de gardons et Y le nombre de carpes obtenus sur les 8 poissons pêchés. Quelles sont les lois des variables aléatoires X et Y? Quelle est la probabilité que sur les 8 poissons pêchés, il y ait :

- Au moins 5 gardons.
- Au plus 4 carpes.

- On pêche 100 poissons. Quelle est la probabilité que le nombre de gardons pêchés soit plus grand que 47?

- Sur 100 poissons pêchés on a effectivement observé 47 gardons, que peut-on en conclure quant à la vraie valeur de p?

1.6 Elections

On se propose de déterminer la taille de l'échantillon à prélever pour que, dans un sondage d'opinion concernant une élection à deux candidats A et B, le candidat gagnant sur l'échantillon soit le candidat gagnant au niveau de la population. (On rappelle que, d'après la théorie de l'échantillonnage, ceci ne peut être obtenu qu'avec une "forte probabilité"). Si p désigne la proportion exacte d'électeurs ayant l'intention de voter A et $\hat{p}$ la proportion observée correspondante, il s'agit donc de prélever un échantillon d'effectif n tel que $\hat{p} > 0,5$ si $p > 0,5$ (et $\hat{p} < 0,5$ si $p < 0,5$) avec "forte probabilité".

- D'une façon générale (c'est-à-dire sans préciser la valeur de n , ni celle de p):
  - Quelle est la distribution d'échantillonage de $\hat{p}$ (moyenne, écart- type, forme)?
  - Quelle est la probabilité pour que $\hat{p}$ soit inférieure à $0,5$?

- On donnera pour chacune des cases du tableau suivant la probabilité pour que $\hat{p}$ soit inférieure à $0,5$.

n p	0,4	0,45	0,475
100			
400			
1000			

- Les résultats obtenus ci-dessus confirment qu'on peut ne pas être certain que p est $< 0,5$, si $\hat{p} < 0,5$. On se propose donc de vouloir limiter à 5% le risque que $\hat{p} > 0,5$, quand $p = 0,48$. Combien faut-il interroger d'électeurs?

1.7 Une souris verte..

On inocule une maladie toujours mortelle sans traitement à des souris afin de savoir si un certain produit peut enrayer cette maladie. On observe un échantillon de 100 souris et on compte le nombre le nombre de souris guéries. Sur cet échantillon, on dénombre 64 guérisons.

- Donner un intervalle de confiance pour p proportion de guérison.

- Quelle aurait dû être la taille minimum n de l'échantillon pour déterminer cette proportion à 2% près, avec le coefficient de confiance 99%?

1.8 Grippe

On effectue une enquête, durant une épidémie de grippe, dans le but de connaître la proportion p de personnes présentant ensuite des complications graves. On observe un échantillon représentatif de 400 personnes et pour un tel échantillon 40 personnes ont présenté des complications.

- Donner un intervalle de confiance pour p au risque 5%.
- On désire que la valeur estimée $\hat{p}$ diffère de la proportion inconnue exacte p de moins de $0,005$ avec une probabilité égale à 95%. Quel sera l'effectif d'un tel échantillon?
- Quel devrait être le risque pour obtenir le même intervalle qu'à la question précédente en conservant l'effectif $n = 400$. Quelles conclusions peut-on en tirer?
- Si on dispose de 40 échantillons, combien en moyenne d'intervalles de confiance (au risque 5%) comprendront la valeur exacte inconnue p?

1.9 Hérédité

On dit qu'un caractère phénotypique est déterminé génétiquement par un couple d'allèles (A,a) ayant un rapport de dominance simple, si la correspondance entre le génotype et le phénotype est le suivant :

génotype	phénotype
aa	(a)
Aa	(A)
AA	(A)

- On croise des individus de phénotype (a) et (A) (parents: génération F_0) puis leurs descendants de phénotype (A) entre eux (enfants : génération F_1). Montrez qu'à la seconde génération (petits-enfants: génération F_2) on s'attend à trouver un quart de phénotype (a) et trois quarts de phénotype (A).
- Inversement, on peut comparer la proportion de phénotype (a) dans la génération F_2 à la valeur théorique $p = 0,25$ pour tester, par une expérience de ce type, le fait qu'un caractère est déterminé par deux allèles à rapport de dominance simple.
  - Sur 40 petits enfants on a compté 8 (a) et 32 (A), montrez que ce résultat n'est pas contraire au schéma (2 allèles dominance simple) au risque d'erreur 5%.
  - On peut en fait se demander si l'examen de 40 petits enfants est suffisant pour mettre en évidence un écart éventuel de p la proportion exacte du phénotype (a) à la valeur théorique $0,25$; en termes statistiques on peut s'interroger sur la puissance du test qu'on vient de faire.
    - ∗ D'une façon générale (c'est-à-dire sans préciser ni la valeur de p, ni le nombre n de petits enfants).
      - · Quelle est la distribution d'échantillonnage de la proportion observée $\hat{p}$ de phénotype (a)?

· Quelle est l'expression du manque de puissance du test de $p = 0,25$ au risque d'erreur 5%?

1.10 Cinéma

A la sortie de 2 salles de cinéma donnant le même film, on a interrogé des spectateurs quant à leur opinion sur le film. Les résultats de ce sondage d'opinion sont les suivants:

Opinion	Mauvais film	Bon film	Total
salle 1	30	70	100
salle 2	48	52	100
Total	78	122	200

- Pour savoir si l'opinion est liée à la salle, quelles proportions observées faut-il comparer? Montrez que l'opinion est significativement liée à la salle, au risque d'erreur 5%.
- On a en outre demandé aux spectateurs interrogés leur âge et on les a classés en 2 catégories: âge < 25 ans, âge > 25 ans.

Opinion	Mauvais film	Bon film	Total
Salle1 âge < 25	18	62	80
Salle 1 âge > 25	12	8	20
Salle 2 âge < 25	7	23	30
Salle 2 âge > 25	41	29	70

- La distribution des opinions est-elle significativement liée à l'âge du spectateur?
- Le choix de la salle de cinéma est-il significativement liée à l'âge du spectateur?
- Si on tient compte de l'âge du spectateur, la distribution des opinions diffère-t-elle significativement dans les 2 salles?

2 Statistique sur les paramètres d'une loi normale

2.1 Inflation

Le gouvernement d'un pays a décidé de fixer, à l'échelon national, le prix d'un produit. Il tolère une distribution des prix suivant une loi normale de moyenne 100 francs et d'écart type 10 francs. Ne pouvant vérifier les nombreux points de vente, il considère un échantillon de 36 points de vente où la moyenne de prix du produit vaut 105 francs. Doit-il considérer que ces pris sont en dehors de la norme imposée?

2.2 Instrument

L'erreur faite quand on utilise un certain instrument de mesure de longueur est une loi normale centrée d'écart-type 1mm.

a) Calculer la probabilité que l'erreur faite quand on utilise une fois l'instrument soit inférieure a 0.5 mm.

b) On suppose que l'instrument est utilisé de façon indépendante 9 fois pour mesurer une longueur donnée. Calculer la probabilité que la moyenne de ces 9 mesures soit au plus distante de 0.5 mm de la vraie longueur.

c) Un autre instrument a été utilisé 10 fois de façon indépendante pour mesurer une longueur donnée. Les erreurs observées sont (en mm): $-0.2, 0.1, -0.3, 0.1, 0.2, -0.3, -0.1, 0.1, 0.2,$ En supposant que ces erreurs sont normalement distribuées, calculer un intervalle de confiance au niveau 95%, pour l'erreur moyenne.

2.3 Heures de vol

Le temps de vol d'un certain type d'avion sur un trajet fixé possède une moyenne de $16,25$ heures depuis la mise en service de ces avions. La distribution du temps de vol a un écart type de $1,5$ heures. Des mesures récentes couvrant 120 vols donnent une moyenne de 15 heures 56 minutes.

a) Donner un intervalle de confiance pour le temps moyen de vol (risque 5%).

b) Y-a-t-il une différence significative avec la moyenne annoncée?

2.4 Course et altitude

Des athlètes ont réalisé une course de 400 mètres au niveau de la mer et en haute altitude leurs temps sont les suivants (en seconde):

Coureur	A	B	C	D	E	F	G	H	I	J
Niveau mer	48.3	47.6	49.2	50.3	48.8	51.1	49.0	48.1	50.7	47.9
Haute alt	50.4	47.3	50.8	52.3	47.7	54.5	48.9	49.9	54.8	48.5

Tester l'hypothèse: la performance des athlètes n'est pas affectée par l'altitude.

2.5 Câblé

Une compagnie fabrique des câbles d'acier dans 2 usines X et Y. Deux échantillons de morceaux de câbles choisis au hasard, d'une longueur de 10 mètres, ont été extraits respectivement des usines X et Y: le premier au nombre de 9 et le second au nombre de 16. La charge de coupure de ces câbles $x_i, i = 1..9, y_j, j = 1..16$ a été déterminée en k-Newtons. Les résultats sont:

$$\overline{x} = \frac{1}{9}\sum_{i=1}^{9} x_i = 30.11, \qquad \sum_{i=1}^{9}(x_i - \overline{x})^2 = 0.8013 \qquad \overline{y} = \frac{1}{16}\sum_{j=1}^{16} y_j = 29.63, \qquad \sum_{j=1}^{16}(y_j - \overline{y})^2 = 3.0206.$$

Y-a-t-il évidence que la charge de coupure est différente pour les câbles des 2 usines X et Y?

2.6 Tiges

Un technicien d'atelier veut vérifier le fonctionnement de deux machines A et B qui doivent produire des tiges d'une longueur identique. Toutes les deux, sur une longue période passée, ont donné des tiges dont la longueur est à peu près une variable normale de variance $0.2cm^2$. En fait, le fonctionnement des machines est maintenant suspect. Pouvez-vous examiner cette suspicion à sa place si la longueur totale de 20 tiges produites par la machine A vaut $824cm$ et celle de 10 tiges par la machine B $395cm$?

2.7 La course fantastique

On admet que le temps d'une course de 100 m pour un individu de sexe masculin choisi au hasard dans une classe de terminale d'un lycée est une variable aléatoire normale de moyenne 12.5 s et d'écart type 0.7 s.

a) On observe 10 individus. Soit X le nombre d'individus parmi les 10 qui font un temps inférieur à 12s. Quelle est la loi de X? Quelle est la probabilité que X soit plus grande que 7?

b) On observe effectivement sur 10 individus, 7 individus qui font un temps inférieur à 12 s, que peut-on en conclure?

2.8 Lapins

11 lapins ont été examinés pour contrôler le taux d'acide urique dans leur plasma. On a obtenu les résultats suivants (en mg/100 ml): $160, 168, 154, 156, 172, 163, 169, 175, 150, 167, 166$. Donner un intervalle de confiance au niveau 90% pour la moyenne de la population de laquelle l'échantillon a été extrait. Un expérimentateur voudrait estimer la vraie valeur du taux d'acide urique à 1% près avec un risque de 10%. Combien d'observations nécessite-t-il?

2.9 Porcs

Pour comparer 2 régimes A et B pour des porcs, 8 paires de porcs ont été utilisées. Les 2 porcs d'une paire sont issus d'une même portée. On a assigné au hasard sur chaque élément de chaque paire le régime A ou B. Les gains de poids pour la durée de l'expérience sont les suivants:

Paire	1	2	3	4	5	6	7	8
Reg A	25.6	20.7	14.0	21.5	21.6	25.6	26.6	22.2
Reg B	24.1	17.7	14.3	19.7	22.2	23.5	25.4	21.3

Y-a-t-il une différence significative au niveau du gain de poids entre les 2 régimes?