

1^{er} baccalauréat en Sciences Vétérinaires

Biostatistiques – Travaux dirigés

Séance de TD n°3

1. Pour une distribution de chi-carré à 12 ddl, trouver la valeur de chi-carré telle que
 - a) L'aire à droite de cette valeur soit égale à 0,05
 - b) L'aire à gauche de cette valeur soit égale à 0,99
 - c) L'aire à droite de cette valeur soit égale à 0,025
2. On procède à 200 parties de pile ou face avec une pièce de monnaie. On observe 115 faces et 85 piles. Tester si la pièce est parfaite au seuil 5%.
3. Afin de tester l'efficacité d'un régime, un expérimentateur souhaite connaître la distribution (à priori, c'est-à-dire sans régime) du "gain quotidien moyen (GQM)" dans une population bovine. Il fait les mesures sur 5 groupes de 20 animaux: le gain moyen (pour les 5 groupes) est de 1400 grammes, et la déviation standard de la moyenne pour les cinq mesures est de 100 grammes (erreur standard). Quelle proportion de groupes de 20 bovins de cette population aura-t-elle un GQM > 2000 grammes (lire: dans quel pourcentage de cas la valeur moyenne dépasse-t-elle 2000 g) ? Quelle est la déviation standard dans la population bovine (données individuelles) ? En supposant la distribution des GQM individuels normale, quelle proportion de bovins aura un GQM > 2000 grammes ?
4. Un chercheur désire savoir si une nouvelle molécule a un effet sur la pression sanguine systolique. Si l'effet sur la moyenne de la population est de 10 mm Hg (faisant par exemple passer la moyenne de 100 mm Hg à 90 mm Hg), quelle devra être la taille de l'échantillon à utiliser pour permettre de détecter cet effet avec une puissance de 80% ?
5. Un nombre est tiré au hasard entre 0 et 10, et chaque valeur est à priori équiprobable (distribution uniforme). Si on tire des ensembles de 9 nombres et qu'on en calcule la moyenne, donnez la distribution de ces moyennes. (Rappel: $\mu = E[x]$, et $\sigma^2 = E[(x-\mu)^2]$)
6. Sur 50 veaux culards du sexe mâle, 34 sont nés par césarienne; sur 40 veaux culards du sexe femelle, 11 sont nés par césarienne. Le sexe influence-t-il les circonstances du vêlage ? Comment exprimez-vous votre H0? Dans votre conclusion, quel type d'erreur commettez-vous et avec quelle probabilité ?
7. L'examen à l'abattage de 1000 porcs a révélé que 300 avaient des lésions pulmonaires et 100 des lésions de la plèvre. 620 étaient absolument indemnes de toute lésion. Y a-t-il indépendance entre les lésions du poumon et les lésions de la plèvre ? Ou encore, un individu ayant des lésions pulmonaires a-t-il plus de chance d'avoir en même temps des lésions de la plèvre qu'un individu n'en ayant pas ?
8. On a comparé des mères primipares et des mères multipares pour la répartition des poids à la naissance, dans 3 classes :

Classes de poids	Primipares	Multipares	Totaux
< 3 kg	26	20	46
3-4 kg	61	63	124
> 4 kg	8	22	30
Totaux	95	105	200

Testez l'indépendance entre ces deux variables, parité et poids à la naissance ? Justifiez la validité du test choisi. Formulez concrètement votre conclusion.

9. Trois préparations d'un même vaccin contre la maladie de Marek ont assuré la protection suivante :

- La 1ère: 16 volailles atteintes sur 54 vaccinées
- La 2ème: 25 volailles atteintes sur 55 vaccinées
- La 3ème: 4 volailles atteintes sur 38 vaccinées

Ces 3 préparations ont-elles des pouvoirs protecteurs différents ?

10. Cent patients ayant été victimes d'un infarctus du myocarde ont été suivis pendant 8 ans. La moitié d'entre eux ont été soumis à un régime strict (pauvre en matières grasses, riche en protéines et en vitamines), tandis que l'autre moitié gardait toute liberté dans le choix de ses aliments. Après 8 ans, on a observé 28 survivants dans le groupe à régime strict et 12 survivants dans le groupe à régime libre. Un régime strict est-il indiqué après infarctus du myocarde ?

11. Cuénot a obtenu une F2 de monohybridisme composée de 198 souris grises et 72 souris albinos, le croisement parental étant souris grise X souris albinos. Testez l'hypothèse 3/4 : 1/4 (le caractère blanc est récessif).

12. Dans une région exposée à une maladie, une substance susceptible de protéger les individus contre la pathologie est testée et les résultats obtenus sont reportés dans la table suivante:

	Atteints	Non atteints	
Protégés	2	12	14
Non protégés	8	6	14
	10	18	28

Si la substance n'a aucun effet, on attendrait des proportions identiques d'atteints dans les deux groupes, "protégés" et "non protégés", à savoir 5 sur 14. Bien entendu, des fluctuations aléatoires d'échantillonnage pourraient faire dévier les effectifs réellement observés des valeurs moyennes attendues. On demande :

- a) De calculer la probabilité de ce qui a été observé si l'hypothèse nulle (pas d'effet de la substance) est vraie, ce qui revient à calculer la probabilité, quand on prélève 10 individus dans un groupe total de 28, que 2 appartiennent à un sous-groupe de 14 et les 8 autres à un autre sous-groupe de 14.
- b) De calculer les probabilités des situations qui s'écartent encore plus de la situation correspondant à l'hypothèse nulle (c'est à dire où le nombre d'atteints parmi les protégés est <2, le nombre total d'individus(28), de protégés(14), de non protégés(14), d'atteints(10) et d'indemnes (18) restant inchangés).
- c) A quoi correspond la somme des probabilités calculées dans les deux points précédents et qu'en concluez-vous ?

13. Un généticien désire déterminer si un caractère a un déterminisme de type dominant. Pour réaliser son expérience, il effectue 100 croisements d'individus exhibant le caractère, et obtient 70 individus eux-mêmes porteurs du caractère. Expliquez la démarche, et les conclusions de l'expérience.
14. Une expérience portant sur l'effet d'une nouvelle molécule sensée augmenter la protection contre une pathologie a été menée. 50 individus ont reçu le vaccin, 50 ont reçu un placebo. 20 individus ont manifesté des signes de la maladie, parmi lesquels 8 avaient reçu la molécule. La molécule a-t-elle réellement un effet ?
15. Certains chercheurs ont émis l'hypothèse selon laquelle le développement du cancer du sein pourrait être lié à l'âge lors de la première naissance. Les données suivantes ont été collectées :

Cas/Contrôle	Age à la première naissance					
	< 20	20-24	25-29	30-34	>34	
Cas	320	1206	1011	463	220	3220
Contrôle	1422	4432	2893	1092	406	10245
	1742	5638	3904	1555	626	13465

Y a-t-il réellement un effet de l'âge ? Si oui, dans quel sens s'exerce-t-il ?

16. Dans un sondage sur 1000 personnes, combien de personnes doivent se déclarer "pour" pour pouvoir conclure à une majorité significative de "pour" (au seuil 5%). Effectuez le calcul avec trois méthodes.
17. Dans une expérience d'inoculation d'un agent pathogène, 3 animaux ayant absorbé une substance protectrice n'ont montré aucun signe de maladie au terme de l'expérience, alors que deux ont été atteints. Parallèlement, 4 autres animaux exposés sans protection sont tous atteints. La protection est-elle réellement efficace ?
18. Un élevage de 50 chiens est visité par un vétérinaire, qui y détecte 10 cas d'une affection. Le vétérinaire conclut à un problème alimentaire, modifie les rations et revisite l'élevage deux mois plus tard. Dans les dix chiens atteints, 6 sont rétablis. Par contre, 2 nouveaux cas sont apparus. L'alimentation a-t-elle un effet significatif ?
19. Une maladie a une prévalence de 10% dans une population donnée. Un échantillon de 200 individus est constitué dans une sous-population, et les animaux sont soumis à un test de détection de la maladie. Le résultat du test est que 32 animaux sont atteints. Peut-on conclure que la sous-population considérée est plus exposée à la maladie ? Comment exprimez-vous votre hypothèse nulle ? Expliquez les types d'erreurs que vous pouvez commettre dans vos conclusions, et donnez leur probabilité.
20. Dans la population Blanc-Bleu Belge, la couleur de la robe est donnée essentiellement par le génotype au niveau d'un gène. Deux variants (allèles) sont possibles au niveau de ce gène: N, qui a une fréquence de 0,7 dans la population et n qui a une fréquence de 0,3. Il y donc trois génotypes possibles, qui sont NN, Nn et nn, donnant les robes blanche, bleue et noire respectivement. Dans un échantillon de 1000 individus, 80 sont noirs, 502 sont blancs et 418 sont bleus. Montrez que ces observations sont en accord avec une situation d'indépendance des allèles (situation connue sous le nom d'équilibre de Hardy-Weinberg).

21. On estime la proportion d'animaux atteints d'un parasite donné à 20%. Dans une ferme, 5 animaux sont testés, et aucun n'est parasité. Pourquoi le test du χ^2 n'est-il pas adapté dans cette situation pour vérifier si la ferme est plus épargnée que la moyenne de la population? Quelle solution alternative proposez-vous ? Que peut-on conclure de cette expérience ?
22. Un zootechnicien a réalisé une expérience de pesée sur un ensemble de 250 bovins d'âges similaires. Il a réparti ses observations par classes de poids. Ses observations se répartissent comme suit:

Classes	Fréquence
moins de 180 kg	2
Entre 180 kg et 200 kg	13
Entre 200 kg et 220 kg	26
Entre 220 kg et 240 kg	51
Entre 240 kg et 260 kg	71
Entre 260 kg et 280 kg	60
Entre 280 kg et 300 kg	23
Entre 300 kg et 320 kg	3

Il aimerait s'assurer que la distribution de poids est gaussienne. Expliquez la démarche.

23. Un généticien prétend que les individus porteurs d'un génotype particulier au niveau d'un gène qu'il a identifié ont tendance à être protégés contre une maladie. Pour démontrer ses dires, il collectionne un échantillon de 100 personnes porteuses du génotype et 100 personnes porteuses d'autres génotypes. Parmi ces dernières, 30 sont atteintes. Combien peuvent être atteintes dans son échantillon de porteurs pour que l'effet soit démontré (au seuil $\alpha = 5\%$) ? Comment s'appelle ce type d'expérience ? Quelles proportions n'ont pas de sens biologique particulier ?
24. Une firme pharmaceutique souhaite commercialiser un médicament utile pour traiter l'arthrite chez les chiens âgés. Pour cela, elle envoie à 100 vétérinaires deux colis à un mois d'intervalle. Un des colis contient effectivement le médicament, et l'autre un *placebo* (évidemment, les destinataires ne sont pas prévenus qu'une des substances servira de témoin). La firme demande ensuite aux vétérinaires quelle substance ils jugent plus efficace. 16 d'entre eux les ont trouvées efficaces toutes les deux. 38 les ont jugées inopérantes. Finalement, 27 ont préféré le traitement 1 (correspondant au médicament), et les autres ont eu l'impression que le traitement 2 (correspondant au *placebo*) avait un effet plus favorable. Comment analysez-vous ces données ? Concluez sur l'efficacité du traitement proposé.
25. Un petit test a été réalisé pour tester l'efficacité d'un vaccin: 5 individus ont été vaccinés et 5 autres pas. Les résultats du test sont dans la table suivante :

	Vaccinés	Non Vaccinés	
Malades	1	3	4
Sains	4	2	6
	5	5	10

De quel genre d'analyse s'agit-il ? Pourquoi n'est on pas intéressé par toutes les conclusions que l'on pourrait tirer de cette analyse ? Comment tester l'efficacité du vaccin ?

26. Générez une table de contingence 3x2 et testez-y l'indépendance. Commentez.

27. On a développé une race de poule « Supercoq » dont la coquille des œufs serait plus solide et on a comparé cette solidité avec celle d'une race traditionnelle :

	Solidité +	Solidité -	
Poules « Supercoq »	209	42	251
Poules traditionnelles	54	95	149
	263	137	400

La race « Supercoq » possède-t-elle des œufs avec des coquilles plus solides ?
Calculez l'Odds Ratio et interprétez-le. Pouvez-vous interpréter le rapport 209/263 ?

28. Quelle est la distance entre les moyennes de deux distributions normales de même variance 10^2 sachant que l'hypothèse nulle d'égalité des moyennes est rejetée 95 fois sur 100 au seuil $\alpha = 5\%$ bilatéral ?
29. Une équipe scientifique a étudié les taux sanguins de cholestérol chez des chiennes hypothyroïdiennes. L'étude a révélé que le taux de cholestérol était normalement distribué avec une moyenne de 208 mg/dl et une déviation standard de 25 mg/dl. Une étude précédente avait montré que des taux inférieurs à 200 mg/dl pouvaient quasiment exclure une hypothyroïdie.
- Quelle est la probabilité d'observer un chien hypothyroïdien avec un cholestérol en dessous de 200 mg/dl ?
 - Pour trois chiens hypothyroïdiens pris au hasard, quelle est la probabilité qu'au moins un des chiens ait un taux de cholestérol sanguin inférieur à 200 mg/dl ?
30. Un éleveur de chiens, dont l'élevage est infecté par l'herpès virose canine, souhaite savoir si la vaccination de ses nouvelles chiennes a été efficace. Pour cela, il regarde la mortalité des chiots dans chacune des portées. Sur 11 portées en 2 ans, 6 ont été observées chez les nouvelles chiennes vaccinées et 1 a perdu sa portée d'herpès virose. Chez les « anciennes » chiennes de son élevage, pas toutes infectées, il a perdu 3 portées d'herpès virose. Quelle est la probabilité de cette situation si on suppose que la vaccination n'a pas eu d'effet (H_0) ?
L'hypothèse nulle est-elle acceptable ?
31. Une étude scientifique analysant l'apparition d'une fièvre de lait après le vêlage dans l'espèce bovine a permis de constater que sur un total de 450 vaches, 89 avaient présenté une fièvre de lait lors de leurs deux premières lactations, 152 n'avaient jamais présenté de fièvre de lait et 113 ont présenté une fièvre de lait uniquement lors de la première lactation. Énoncez l'hypothèse nulle et testez si le numéro de lactation a un impact sur l'incidence d'apparition de la fièvre de lait. Interprétez vos résultats.
32. Chez les Yorkshire Terriers, la mise-bas pose souvent des problèmes en raison de la conformation de la race. Dans une étude effectuée sur 160 chiennes, on a essayé de voir si le fait de présenter une dystocie (accouchement difficile) lors de la première portée était susceptible de se représenter lors de la seconde ou non. Les résultats sont donnés dans la table suivante:

		Nichée 2		Totaux
		Dystocie	Pas de dystocie	
Nichée 1	Dystocie	21	31	52
	Pas de dystocie	37	71	108
Totaux		58	102	160

- a) Quelle hypothèse (nulle) veut-on tester dans ce test ?
Quelle procédure utiliser pour mettre en évidence l'éventuelle relation ?
- b) Si l'hypothèse nulle est vraie, combien de chiennes, non dystociques en première portée, *attend-on* devenir dystociques en deuxième portée ?
- c) Combien vaut la statistique utilisée pour tester cet effet ?
- d) Est-ce que le test est significatif (càd. permet de rejeter H_0) ?

33. Un référendum est effectué sur 500 personnes. Déterminez de deux manières la majorité significative au seuil 5%.

34. On souhaite tester l'efficacité d'une vaccination. Pour cela, 300 chiens sont choisis dont 2/3 ont été vaccinés et on constate que 18 chiens vaccinés sont tombés malades contre 32 chez les non vaccinés après exposition au virus. La vaccination est-elle efficace ? Calculez l'Odds Ratio et interprétez-le. Sachant que sur les 50 chiens tombés malades, 18 étaient vaccinés, pouvez-vous interpréter la probabilité de 18/50 (36%) ?

35. On a développé une race de poule « Supercoq » dont la coquille des œufs serait plus solide et on a comparé cette solidité avec celle d'une race traditionnelle :

	Solidité +	Solidité -	
Poules « Supercoq »	7	2	9
Poules traditionnelles	2	5	7
	9	7	16

- a) Quelle est la probabilité de cette situation ?
- b) L'hypothèse nulle voulant une égalité de « solidité + » entre les deux races est-elle vraisemblable ? Interprétez la zone de rejet.