

## Probabilités et distributions continues

### Exercice 1

Si, dans une race bovine, la durée moyenne de gestation est de 280 jours avec déviation standard égale à 5, quelle est la proportion des vaches dont la durée de gestation s'écarte de +/- 10 jours de la moyenne?

moyenne = 280  
dév. stand. = 5

En utilisant la formule suivante:

$$Z = \frac{X - \mu}{\sigma}$$

et sachant que  $X_1=270$  et  $X_2=290$  puisque que dans l'énoncé, on parle de durée de gestation s'écartant de +/- 10

jours par rapport à la moyenne de 280 jours! On peut aisément trouver les valeurs de Z suivantes:

$Z_1 = -2$   
 $Z_2 = 2$

En regardant dans une table des surfaces, on constate que pour une valeur de  $Z=2$  correspond une probabilité de 0,4772, de même que pour le  $Z=-2$ . Donc la probabilité d'être compris dans l'intervalle +/- 10 jours autour de la moyenne de 280 jours est de  $2 \times 0,4772 = 0,9544 = 95,44 \%$ .

Et donc,  $100 - 95,44 = 4,56 \%$  est la proportion de vaches dont la durée de gestation s'écarte de +/- 10 jours de la moyenne.

### Exercice 2

Sachant que, en moyenne, 50 % des œufs donnent des poussins mâles, combien d'œufs faut-il mettre en incubation pour avoir au moins 20 poulettes, dans 99,9 % des cas?

Ici, on veut connaître la valeur de n pour avoir au moins 20 poulettes, il s'agit donc d'un test unilatéral. Il faut donc regarder dans la table des surfaces, quelle est la valeur de Z associée à une probabilité de 0,499. C'est à dire un Z de -3,1 avec un - puisqu'il s'agit d'un test unilatéral gauche!

Ensuite, il suffit de remplacer les différentes données dans la formule suivante:

$$Z = \frac{r - np}{\sqrt{npq}}$$

Ce qui donne: 
$$-3.1 = \frac{20 - 0.5n}{\sqrt{0.5^2 n}} = 0.25 n^2 - 24.805 n + 400 = 0$$

Pour rappel, les formules à utiliser pour résoudre une équation du second degré à une inconnue sont:

$$\Delta = b^2 - 4ac$$

$$n = \frac{-b \pm \sqrt{\Delta}}{2a}$$

Delta= 215,288025  
n1= 78,95539316  
n2= 20,26460684

La valeur de n2 étant absurde vu la proba de 50 % et le fait de vouloir au moins 20 poulettes, on peut dire qu'il faut mettre 79 œufs en incubation pour avoir au moins 20 poulettes!

### Exercice 3

Quelle proportion de l'aire délimitée par la courbe de Gauss et l'axe des X, est comprise entre les limites suivantes:

- a) de - l'infini à -2,937.
- b) de -0,54 à 1,235.
- c) de +1,839 à +2,312.
- d) de +2,51 à + l'infini.

a) En regardant dans la table des surfaces, on se rend compte que la probabilité d'avoir un Z compris entre -2,937 et 0 est de 0,4984 et donc la proportion de l'aire comprise entre - l'infini et -2,937 est de: 50-49,84= 0,16 %

b) Proba. de Z compris entre -0,54 et 0 est de 0,2054 et la proba. de Z compris entre 0 et 1,235 est de 0,3916. Donc la proportion d'aire comprise entre -0,54 et 1,235 est de 20,54+39,16 = 59,7 %

c) Proba. de Z compris entre 0 et 1,839 est de 0,46703 et la proba. de Z compris entre 0 et 2,312 est de 0,48964. Donc la proportion d'aire comprise entre 1,839 et 2,312 est de 48,964-46,703 = 2,261 %

d) Proba. de Z compris entre 0 et 2,51 est de 0,494 et la proba. de Z compris entre 0 et + l'infini est bien entendu de 0,5. Donc la proportion d'aire comprise entre 2,51 et +l'infini est de 50-49,4 = 0,6 %

#### Exercice 4

a) On a calculé les numérations globulaires ( en 10 exp.6/mm³) de 20 porcelets âgés de 20 semaines:

|     |     |     |     |
|-----|-----|-----|-----|
| 4,3 | 6,4 | 7,4 | 9,1 |
| 5,2 | 6,6 | 7,5 | 9,3 |
| 5,3 | 6,7 | 7,8 |     |
| 5,8 | 6,9 | 8   |     |
| 6   | 7,1 | 8,4 |     |
| 6,3 | 7,2 | 8,7 |     |

Ces données sont distribuées normalement avec moyenne de 7 et déviation standard de 1,2. On demande quel pourcentage de porcelets aura une numération supérieure à 9?

b) Un porcelet, ayant reçu une alimentation spéciale, a une numération de 9,2. Quelle est la probabilité qu'une numération aussi élevée soit due uniquement à la chance?

a) Moyenne= 7  
Dév.st.= 1,2  
n= 20

En utilisant:  $Z = \frac{X - \mu}{\sigma}$  et en remplaçant X par 9, on obtient Z= 1,66666667

Cette valeur de Z correspond à une probabilité de 0,4521 et donc le pourcentage de porcelets ayant une numération supérieure à 9000000/mm³ est de 50-45,21= 4,79 %

b) Z= 1,833333333  
Proba= 0,4664  
Et donc, la probabilité d'une numération de 9,2 due uniquement à la chance= 0,0336

#### Exercice 5

On considère généralement que le poids des enfants à la naissance suit une distribution

normale, de moyenne 3,4 kg et de variance 0,25 kg<sup>2</sup>. On suspecte cependant que les femmes diabétiques mettent au monde des enfants qui ont en moyenne un poids inférieur à 3,4 kg. Afin de vérifier cette hypothèse, on a relevé le poids de 25 enfants nés de mères diabétiques et le poids moyen observé a été de 3,74 kg. On demande:

a) Quelle est la probabilité d'observer un poids moyen aussi élevé ou plus élevé si les enfants nés de mères diabétiques obéissent à la loi générale?

b) Quelle hypothèse acceptera-t-on?

a) Moy.= 3,4  
 Var.= 0,25  
 Dév.st.= 0,5  
 n= 25  
 x moy.= 3,74

Attention, il faut bien faire la différence entre les deux formules suivantes:

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$Z = \frac{X - \mu}{\sigma}$$

La première est à utiliser lorsqu'on est en présence d'un individu possédant une valeur X, à comparer avec la moyenne de population. Tandis que la deuxième s'utilise quand on est en présence d'un échantillon d'individus (25 dans le cas présent!) à comparer avec la moyenne de population! Ce qui nous donne un Z = 3,4 correspondant à une probabilité de 0,4997 et donc la probabilité d'observer un poids moyen de 3,74 ou supérieur est de 0,5-0,4997= 0,0003

b) Hypothèse: les femmes diabétiques mettent au monde des enfants qui ont en moyenne un poids à 3,4 kg. La probabilité d'observer un poids moyen (de 3,74 kg) aussi élevé et plus élevé, si les enfants obéissent à la loi générale est de 0,03 % < 0,1 %. Donc rejet au seuil 0,1 %. Donc, ces enfants n'obéissent pas à loi générale!

**Remarque:** Si on avait pris un individu, né de mère diabétique de 3,74 kg, on aurait utilisé la formule suivante:

$$Z = \frac{X - \mu}{\sigma}$$

et obtenu comme résultat Z =

0,68

ca qui correspond à une probabilité de 0,2517.

En conclusion, nous dirons donc que la probabilité d'observer un poids aussi élevé ou plus élevé pour un individu est plus élevé que la probabilité d'observer un poids moyen (d'un échantillon) aussi élevé ou plus élevé.

## Exercice 6

Dans une population humaine, la quantité d'urée dans le sang, exprimée en mg %, est une variable aléatoire normale de moyenne 27 et d'écart-type 5 chez les hommes, tandis que chez les femmes, la moyenne est de 26,4 et la variance 64.

a) Lorsqu'on prend un individu au hasard, calculez la probabilité que son urémie soit inférieure à 22,4 mg % suivant qu'il s'agit d'un homme ou d'une femme.

b) Déterminez une limite  $x_1$ , appelée seuil pathologique, telle qu'il y ait 2,5 % d'hommes dont l'urémie est supérieure à  $X_1$ .

a) Homme:    moyenne=                    27  
                   dév.st.=                    5  
                    $X=$                             22,4  
                    $Z=$                             -0,92  
                   Proba=                    0,17878635

Femme:    moyenne=                    26,4  
                   dév.st.=                    8  
                    $X=$                             22,4  
                    $Z=$                             -0,5  
                   Proba=                    0,30853753

b) Dans la table, 0,475 correspond à un  $Z = 1,96$ ; en modifiant la formule comme suit:

$$Z = \frac{X - \mu}{\sigma} \Leftrightarrow X = Z\sigma + \mu$$

cela nous permet de trouver le seuil pathologique qui est de :                    36,8

### Exercice 7

Dans un élevage de souris, la taille des portées est en moyenne de 10 jeunes, avec déviation standard 2,5. Combien faut-il produire de nichées pour avoir 97,5 chances sur 100 de produire au moins 50 jeunes? On admet que la taille des portées est distribuée normalement.

La formule à utiliser ici est la même que celle de l'exercice précédent, mais en la modifiant légèrement comme suit:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\frac{X}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{X - n\mu}{\sigma\sqrt{n}}$$

car en effet, nous ne disposons pas de la moyenne et nous recherchons la taille de l'échantillon!

Une façon différente de voir la chose, mais qui mène au même résultat et vous semblera certainement plus compréhensible pour l'obtention de la formule finale ci-dessus, est de partir de la formule générale de la loi normale:                     $Z = \frac{X - \mu}{\sigma}$  où le X est ici une somme de variables

(de portées) égale à 50, la moyenne est ici la moyenne d'une somme de variables continues distribuées normalement et égale à la somme des moyennes et donc =  $n\mu$  et la variance d'une somme de variables continues distribuées normalement est égale à la somme des variances, ajoutée de  $n*(n-1)/2$  termes de covariance entre ces variables (nuls dans ce cas-ci car les variables sont indépendantes) et donc la déviation standard (=racine de la variance) est:  $\sigma\sqrt{n}$   
 Voir cours théorique page 27!

La valeur de Z associée est celle correspondant à une probabilité de 0,025 puisqu'il s'agit d'un test unilatéral gauche (les 97,5 % supérieurs) et cette valeur s'obtient comme suit:                    -1,95996108

$X:$                             50  
 Moyenne:                    10  
 Dév. Stand.:                    2,5

$$-1,96 = \frac{50 - 10n}{\sigma\sqrt{n}}$$

1048596,48

En remplaçant dans la formule ci-dessus, on obtient:

$$-1.96 = \frac{50 - 10n}{2.5\sqrt{n}}$$

Ce qui mène à l'équation à une inconnue, du second degré suivante:

$$24.01n = 100n^2 - 1000n + 2500$$

ou  $100n^2 - 1024.01n + 2500 = 0$

La méthode du delta est à employer ici:

$$\Delta = b^2 - 4ac$$

Soit=

48596,4801

Et la taille de l'échantillon se trouve comme suit:

$$n = \frac{-b \pm \sqrt{\Delta}}{2a}$$

Soit=

6,22228047

Soit=

4,01781953

Evidement, la première solution est la bonne car la deuxième est absurde (pourquoi ?) !

### Exercice 8

A supposer que le taux de fécondation soit de 60 %, combien de vaches faut-il inséminer pour produire au moins 15 femelles, dans 97,5 % des cas? Donnez les formules, indiquez la marche à suivre en remplaçant les variables par leurs valeurs. (Egale proportion des sexes).

Taux de fécondation de 60 %, donc probabilité de 0,6.

On demande de trouver le nombre de vaches à inséminer pour obtenir 15 veaux femelles dans 97,5 % des cas. Il ne s'agit pas d'une variable continue mais discontinue (distribution binomiale) car présence de femelles ou de mâles ( 2 possibilités). Passage à la distribution normale car p proche de 0,5, n est l'inconnue mais doit être grand, nous pouvons utiliser la formule de l'approximation Gaussienne d'une variable binomiale:

$$Z = \frac{X - \mu}{\sigma} = \frac{r - np}{\sqrt{npq}}$$

On doit produire au moins 15 femelles dans 97,5 % des cas, cela implique qu'il faut utiliser la valeur de Z correspondant à 2,5 % à gauche de la courbe (test unilatéral gauche), c'est à dire: -1,96

La probabilité d'avoir un veau femelle (p=0,5) après insémination (p=0,6) est de:

0,3

q= 0,7  
r= 15

$$Z = \frac{r - np}{\sqrt{npq}} \Rightarrow -1.96 = \frac{15 - 0.3n}{\sqrt{(0.3)(0.7)n}}$$

$$0.09n^2 - 9,806736n + 225 = 0$$

Delta= 15,17207097

n1= 76,12150149

n2= 32,84223184

La première des solutions est la bonne, car la deuxième est beaucoup moins crédible!

Pourquoi?

Nous dirons donc qu'il faut inséminer 77 vaches pour obtenir 15 veaux femelles dans 97,5% des cas, dans ce cas-ci (0,6 de proba de fécondation et sex ratio de 0,5)!

### Exercice 9

H0: le taureau est indemne

H1: le taureau est porteur.

Le principe est le suivant: si un descendant est atteint, on suppose que le taureau est porteur et on rejette l'hypothèse nulle.

Si aucun descendant n'est atteint, on suppose que le taureau est sain et on accepte H0

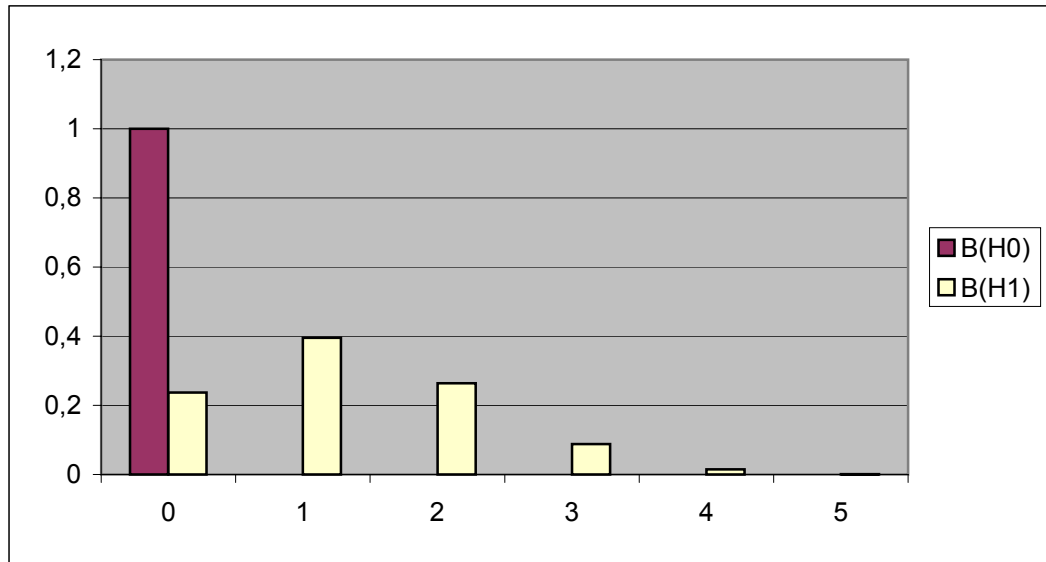
L'événement à considérer est donc l'apparition d'un descendant atteint (binomial)

sous H0 p=0

sous H1 p=0.25

n=5

|   |       |       |            |
|---|-------|-------|------------|
| n | p     | p     |            |
|   | 5     | 0     | 0,25       |
| r | B(H0) | B(H1) |            |
|   | 0     | 1     | 0,23730469 |
|   | 1     | 0     | 0,39550781 |
|   | 2     | 0     | 0,26367188 |
|   | 3     | 0     | 0,08789063 |
|   | 4     | 0     | 0,01464844 |
|   | 5     | 0     | 0,00097656 |



Erreur de type I: Rejeter H0 erronément

$\alpha=0.0$

Erreur de type II: Accepter l'hypothèse nulle erronément.

$\beta=P(r=0|H1)=0.2373$

Puissance

$P=1-\beta=0.7627$

Nombre de croisements = N :  $b = 1 - P(r=0 | H1, N) = 0.99$

| N  | Puissance   |
|----|-------------|
| 5  | 0,762695313 |
| 6  | 0,822021484 |
| 7  | 0,866516113 |
| 8  | 0,899887085 |
| 9  | 0,924915314 |
| 10 | 0,943686485 |
| 11 | 0,957764864 |
| 12 | 0,968323648 |
| 13 | 0,976242736 |
| 14 | 0,982182052 |
| 15 | 0,986636539 |
| 16 | 0,989977404 |
| 17 | 0,992483053 |

#### Exercice 10

$z = (X - \mu) / \sigma$

$z = -1,333333333$

$$P(z > -1.33333) = 0,90878872$$

$$P = 0.95 \Rightarrow z = -1.64 \Rightarrow \mu = X - z\sigma = 5000 + 1.64 * 750 = 6230$$

$$P(5) = P(z > -1.3333) ^ 5 = 0,61989001$$

### Exercice 11

La taille d'une portée  $\sim N(10;2.5)$

La taille de n portées indépendantes est distribuée comme la somme de n variables:

la moyenne est la somme des n moyennes

la variance est la somme des variances (les covariances sont nulles)

la deviation standard est  $\text{racine}(n \cdot \sigma^2) = \sigma \cdot \text{racine}(n)$

$$z = -1.96 = (X - \mu) / \sigma = (50 - 10 \cdot n) / (2.5 \cdot \text{racine}(n))$$

x                      racine(n)

a                              10

b                              -4,9

c                              -50

x1                      2,494449933 n1                      6,22228047

x2                      -2,004449933 n2                      4,01781953 => rejeté

### Exercice 12

1) Approche binomiale

$$p = 0,5$$

$$B(r \geq 20) = 0.95 \Rightarrow B(r < 20) = 0.05$$

| n     |  |             |            |            |            |            |
|-------|--|-------------|------------|------------|------------|------------|
| r     |  | 30          | 40         | 50         | 55         | 51         |
| 0     |  | 9,31323E-10 | 9,0949E-13 | 8,8818E-16 | 2,7756E-17 | 4,4409E-16 |
| 1     |  | 2,79397E-08 | 3,638E-11  | 4,4409E-14 | 1,5266E-15 | 2,2649E-14 |
| 2     |  | 4,05125E-07 | 7,0941E-10 | 1,088E-12  | 4,1217E-14 | 5,6621E-13 |
| 3     |  | 3,78117E-06 | 8,9858E-09 | 1,7408E-11 | 7,2817E-13 | 9,2482E-12 |
| 4     |  | 2,55229E-05 | 8,3119E-08 | 2,0455E-10 | 9,4662E-12 | 1,1098E-10 |
| 5     |  | 0,000132719 | 5,9845E-07 | 1,8818E-09 | 9,6555E-11 | 1,0432E-09 |
| 6     |  | 0,000552996 | 3,491E-06  | 1,4114E-08 | 8,0463E-10 | 7,9978E-09 |
| 7     |  | 0,001895986 | 1,6956E-05 | 8,8715E-08 | 5,6324E-09 | 5,1414E-08 |
| 8     |  | 0,005450961 | 6,9944E-05 | 4,7684E-07 | 3,3794E-08 | 2,8278E-07 |
| 9     |  | 0,013324572 | 0,00024869 | 2,2253E-06 | 1,7648E-07 | 1,3511E-06 |
| 10    |  | 0,027981601 | 0,00077094 | 9,1236E-06 | 8,1181E-07 | 5,6744E-06 |
| 11    |  | 0,050875638 | 0,00210257 | 3,3177E-05 | 3,3211E-06 | 2,115E-05  |
| 12    |  | 0,080553093 | 0,00508121 | 0,00010782 | 1,2177E-05 | 7,0501E-05 |
| 13    |  | 0,111535052 | 0,01094415 | 0,00031518 | 4,0278E-05 | 0,0002115  |
| 14    |  | 0,13543542  | 0,02110658 | 0,00083297 | 0,00012084 | 0,00057408 |
| 15    |  | 0,144464448 | 0,03658474 | 0,00199914 | 0,00033028 | 0,00141606 |
| 16    |  | 0,13543542  | 0,05716365 | 0,00437311 | 0,00082571 | 0,00318613 |
| 17    |  | 0,111535052 | 0,08070163 | 0,00874623 | 0,00189427 | 0,00655967 |
| 18    |  | 0,080553093 | 0,10311875 | 0,01603475 | 0,00399902 | 0,01239049 |
| 19    |  | 0,050875638 | 0,11940065 | 0,0270059  | 0,00778756 | 0,02152033 |
| Somme |  | 0,950631427 | 0,43731466 | 0,05946023 | 0,01501447 | 0,04595727 |

2) Approximation normale

$$z = -1.64 = (X - \mu) / \sigma = (20 - n \cdot p) / \text{racine}(n \cdot p \cdot q)$$

x=racine(n)

|       |         |
|-------|---------|
| a     | 0,5     |
| b     | -0,82   |
| c     | -20     |
| Delta | 40,6724 |

|    |                |      |                     |
|----|----------------|------|---------------------|
| x1 | 7,19749167 =>  | n1 = | 51,8038863          |
| x2 | -5,55749167 => | n2 = | 30,8857137 (rejeté) |

### Exercice 13

Poids nécessaire = 200000 / 1600 = 125 mgr

Poids requis = 0.95 \* 125 = 118,75

Proportion (Poids > 118,75) = P (z > ((118,75-126)/4) = P (z > -1,8125) =

0,96504557