Tests statistiques : présentation

Inférence statistique
BTSA Anabiotec et VO

Yannick Le Bastard

LEGTA de l'Hérault

March 12, 2024


Bienvenue sur les
sentiers
mathématiques !

Un premier exemple

La notion de test statistique repose sur le principe suivant :

Nous nous intéressons à un caractère particulier dans une population de référence (taille ou masse moyenne, pH moyen, prix moyen, variance, proportion, etc.), mais ne pouvant tester toute la population, **nous en prélevons un échantillon**.

Un premier exemple

La notion de test statistique repose sur le principe suivant :

Nous nous intéressons à un caractère particulier dans une population de référence (taille ou masse moyenne, pH moyen, prix moyen, variance, proportion, etc.), mais ne pouvant tester toute la population, **nous en prélevons un échantillon**.

Nous ne parlerons pas pour le moment de la manière dont cet échantillon est constitué.

Un premier exemple

La notion de test statistique repose sur le principe suivant :

Nous nous intéressons à un caractère particulier dans une population de référence (taille ou masse moyenne, pH moyen, prix moyen, variance, proportion, etc.), mais ne pouvant tester toute la population, **nous en prélevons un échantillon**.

Nous ne parlerons pas pour le moment de la manière dont cet échantillon est constitué.

Nous effectuons ensuite deux hypothèses incompatibles sur la population, et au vu de notre échantillon et d'un certain risque d'erreur que nous détaillerons, nous sommes amenés à pencher vers une des deux hypothèses.

Un premier exemple

Imaginons par exemple qu'une bouteille d'eau affiche sur son étiquette une contenance de 1,5 L.

Un premier exemple

Imaginons par exemple qu'une bouteille d'eau affiche sur son étiquette une contenance de 1,5 L.

Il s'agit bien entendu d'une contenance moyenne : certaines bouteilles auront une contenance légèrement inférieure, d'autres légèrement supérieure. Mais si l'indication fournie par le fabricant est exacte, nous nous attendons à ce que la contenance moyenne d'une bouteille ne soit pas **significativement différente** des 1,5 L annoncés (la **norme**).

Un premier exemple

Imaginons par exemple qu'une bouteille d'eau affiche sur son étiquette une contenance de 1,5 L.

Il s'agit bien entendu d'une contenance moyenne : certaines bouteilles auront une contenance légèrement inférieure, d'autres légèrement supérieure. Mais si l'indication fournie par le fabricant est exacte, nous nous attendons à ce que la contenance moyenne d'une bouteille ne soit pas **significativement différente** des 1,5 L annoncés (la **norme**).

Comment procéder sachant qu'on ne peut vérifier toute la production ?

Un second exemple

Un laboratoire pharmaceutique souhaite tester l'efficacité d'un traitement préventif (vaccin) pour une pathologie affectant 10% de la population. Pour cela, il effectue un test sur un échantillon de 500 personnes ayant reçu le vaccin.

Un second exemple

Un laboratoire pharmaceutique souhaite tester l'efficacité d'un traitement préventif (vaccin) pour une pathologie affectant 10% de la population. Pour cela, il effectue un test sur un échantillon de 500 personnes ayant reçu le vaccin.

Parmi les individus de ce groupe, 43 ont contracté la pathologie.
Peut-on conclure à l'efficacité du vaccin ?

Un troisième exemple

On cherche à comparer le rendement moyen d'une céréale, cultivée avec les mêmes pratiques dans des conditions édaphiques similaires, mais dans deux régions différentes.

Un troisième exemple

On cherche à comparer le rendement moyen d'une céréale, cultivée avec les mêmes pratiques dans des conditions édaphiques similaires, mais dans deux régions différentes.

Dans la région 1, sur un échantillon de 20 parcelles, on a obtenu un rendement moyen de 55q/ha et un écart-type de 2,5q/ha.

Dans la région 2, sur un échantillon de 25 parcelles, on a obtenu un rendement moyen de 58q/ha et un écart-type de 3,8q/ha.

Un troisième exemple

On cherche à comparer le rendement moyen d'une céréale, cultivée avec les mêmes pratiques dans des conditions édaphiques similaires, mais dans deux régions différentes.

Dans la région 1, sur un échantillon de 20 parcelles, on a obtenu un rendement moyen de 55q/ha et un écart-type de 2,5q/ha.

Dans la région 2, sur un échantillon de 25 parcelles, on a obtenu un rendement moyen de 58q/ha et un écart-type de 3,8q/ha.

Peut-on conclure que les rendements de cette céréale dans la région 1 sont significativement inférieurs à ceux dans la région 2 ?

Un questionnement imprécis (1)

La réponse à la première question était évidemment de prélever un échantillon de la population (l'ensemble des bouteilles de la production), de tester cet échantillon, et en cas de contenance moyenne "*très proche*" de 1,5 L, nous avons envie d'*extrapoler* en disant que tout est correct pour la production.

Un questionnement imprécis (1)

La réponse à la première question était évidemment de prélever un échantillon de la population (l'ensemble des bouteilles de la production), de tester cet échantillon, et en cas de contenance moyenne "*très proche*" de 1,5 L, nous avons envie d'*extrapoler* en disant que tout est correct pour la production.

Mais que signifie "très proche" ?

Un questionnement imprécis (1)

La réponse à la première question était évidemment de prélever un échantillon de la population (l'ensemble des bouteilles de la production), de tester cet échantillon, et en cas de contenance moyenne "*très proche*" de 1,5 L, nous avons envie d'*extrapoler* en disant que tout est correct pour la production.

Mais que signifie "très proche" ?

Comment extrapoler ? Peut-on vraiment le faire ?

Un questionnement imprécis (1)

La réponse à la première question était évidemment de prélever un échantillon de la population (l'ensemble des bouteilles de la production), de tester cet échantillon, et en cas de contenance moyenne "*très proche*" de 1,5 L, nous avons envie d'*extrapoler* en disant que tout est correct pour la production.

Mais que signifie "très proche" ?

Comment extrapoler ? Peut-on vraiment le faire ?

Et comment évaluer l'erreur que l'on fait ?

Un questionnement imprécis (2)

Afin de mieux comprendre cette notion d'erreur, prenons un premier exemple où il n'est ni question de population, d'échantillon et encore moins d'extrapolation : le tribunal !!!

Un questionnement imprécis (2)

Afin de mieux comprendre cette notion d'erreur, prenons un premier exemple où il n'est ni question de population, d'échantillon et encore moins d'extrapolation : le tribunal !!!

Un individu est accusé d'un crime au tribunal. Nous allons poser deux hypothèses :

Un questionnement imprécis (2)

Afin de mieux comprendre cette notion d'erreur, prenons un premier exemple où il n'est ni question de population, d'échantillon et encore moins d'extrapolation : le tribunal !!!

Un individu est accusé d'un crime au tribunal. Nous allons poser deux hypothèses :

$$\begin{cases} H_0 & : \text{ L'individu est innocent} \\ H_1 & : \text{ L'individu est coupable} \end{cases}$$

Un questionnement imprécis (2)

Afin de mieux comprendre cette notion d'erreur, prenons un premier exemple où il n'est ni question de population, d'échantillon et encore moins d'extrapolation : le tribunal !!!

Un individu est accusé d'un crime au tribunal. Nous allons poser deux hypothèses :

$$\begin{cases} H_0 & : \text{ L'individu est innocent} \\ H_1 & : \text{ L'individu est coupable} \end{cases}$$

Deux situations amènent à une décision correcte du tribunal, et deux autres à une erreur.

Un questionnement imprécis (2)

Ce que nous pouvons résumer dans le tableau suivant :

Réalité / Décision	H_0 vraie	H_1 vraie
H_0 décidée	Bonne décision	Mauvaise décision
H_1 décidée	Mauvaise décision	Bonne décision

Un questionnement imprécis (2)

Ce que nous pouvons résumer dans le tableau suivant :

Réalité / Décision	H_0 vraie	H_1 vraie
H_0 décidée	Bonne décision	Mauvaise décision
H_1 décidée	Mauvaise décision	Bonne décision

Nous constatons qu'il existe deux types d'erreur :

1. Déclarer coupable un innocent
2. Déclarer innocent un coupable

Un questionnement imprécis (2)

Ce que nous pouvons résumer dans le tableau suivant :

Réalité / Décision	H_0 vraie	H_1 vraie
H_0 décidée	Bonne décision	Mauvaise décision
H_1 décidée	Mauvaise décision	Bonne décision

Nous constatons qu'il existe deux types d'erreur :

1. Déclarer coupable un innocent
2. Déclarer innocent un coupable

Mais nous ne pouvons pas les quantifier. Prenons un autre exemple.

Un questionnement qui se précise

Une pandémie sévit dans une certaine région du globe. Un test de dépistage est disponible pour celle-ci, mais n'est pas fiable à 100%. Nous cherchons à déterminer si un individu prélevé au hasard dans la population est sain ou malade à l'aide de ce test.

Un questionnement qui se précise

Une pandémie sévit dans une certaine région du globe. Un test de dépistage est disponible pour celle-ci, mais n'est pas fiable à 100%. Nous cherchons à déterminer si un individu prélevé au hasard dans la population est sain ou malade à l'aide de ce test.

Il semble naturel ici de formuler nos deux hypothèses par :

$$\begin{cases} H_0 & : \text{"L'individu est sain"} \\ H_1 & : \text{"L'individu est malade"} \end{cases}$$

Un questionnement qui se précise

Une pandémie sévit dans une certaine région du globe. Un test de dépistage est disponible pour celle-ci, mais n'est pas fiable à 100%. Nous cherchons à déterminer si un individu prélevé au hasard dans la population est sain ou malade à l'aide de ce test.

Il semble naturel ici de formuler nos deux hypothèses par :

$$\begin{cases} H_0 & : \text{ "L'individu est sain"} \\ H_1 & : \text{ "L'individu est malade"} \end{cases}$$

Ce que nous pouvons résumer dans le tableau suivant :

Décision \ Réalité	H_0 vraie	H_1 vraie
H_0 décidée	Vrai négatif	Faux négatif
H_1 décidée	Faux positif	Vrai positif

Un questionnement qui se précise

1. La **sensibilité** (Se) est la probabilité qu'un test de dépistage réalisé sur une personne malade se révèle positif, autrement dit que le test soit positif sachant que la personne est malade.
2. La **spécificité** (Sp) est la probabilité qu'un test de dépistage réalisé sur une personne saine se révèle négatif, autrement dit que le test soit négatif sachant que la personne n'est pas malade.

Un questionnement qui se précise

1. La **sensibilité** (Se) est la probabilité qu'un test de dépistage réalisé sur une personne malade se révèle positif, autrement dit que le test soit positif sachant que la personne est malade.
2. La **spécificité** (Sp) est la probabilité qu'un test de dépistage réalisé sur une personne saine se révèle négatif, autrement dit que le test soit négatif sachant que la personne n'est pas malade.

De la même manière, on peut définir les deux types d'erreur par :

1. Le risque α, dit aussi **risque de première espèce** est la probabilité qu'un test de dépistage réalisé sur une personne saine se révèle positif, autrement dit que le test soit positif sachant que la personne est saine.
2. Le risque β, dit aussi **risque de seconde espèce** est la probabilité qu'un test de dépistage réalisé sur une personne malade se révèle négatif, autrement dit que le test soit négatif sachant que la personne est malade.

Premières définitions

Plus généralement on a les définitions suivantes.

1. Le risque α, dit aussi **risque de première espèce** est la probabilité de décider H_1 alors que H_0 est vraie : $\alpha = P_{H_0 \text{ vraie}}(\text{Décider } H_1)$.
2. Le risque β, dit aussi **risque de seconde espèce** est la probabilité de décider H_0 alors que H_1 est vraie : $\beta = P_{H_1 \text{ vraie}}(\text{Décider } H_0)$.
3. La **puissance du test** est la probabilité de décider H_1 quand H_1 est vraie : $1 - \beta = P_{H_1 \text{ vraie}}(\text{Décider } H_1)$.

Vers une méthodologie

Reprenons l'exemple 2 : Un laboratoire pharmaceutique souhaite tester l'efficacité d'un traitement préventif (vaccin) pour une pathologie affectant 10% de la population. Pour cela, il vaccine un échantillon de 500 personnes. Parmi les individus de ce groupe, 43 ont contracté la pathologie. Peut-on conclure à l'efficacité du vaccin ?

Vers une méthodologie

Reprenons l'exemple 2 : Un laboratoire pharmaceutique souhaite tester l'efficacité d'un traitement préventif (vaccin) pour une pathologie affectant 10% de la population. Pour cela, il vaccine un échantillon de 500 personnes. Parmi les individus de ce groupe, 43 ont contracté la pathologie. Peut-on conclure à l'efficacité du vaccin ?

L'idée naturelle est de calculer la proportion p d'individus infectés dans notre échantillon malgré leur vaccination : $p = \frac{43}{500} = 0,086$.
$p < 0,1$, donc **il semblerait** que le vaccin ait un effet protecteur. Pour autant, rien n'indique que nous ne soyons pas en présence d'un "échantillon exceptionnel". Un autre échantillon aurait pu donner des résultats très différents.

Vers une méthodologie

Reprenons l'exemple 2 : Un laboratoire pharmaceutique souhaite tester l'efficacité d'un traitement préventif (vaccin) pour une pathologie affectant 10% de la population. Pour cela, il vaccine un échantillon de 500 personnes. Parmi les individus de ce groupe, 43 ont contracté la pathologie. Peut-on conclure à l'efficacité du vaccin ?

L'idée naturelle est de calculer la proportion p d'individus infectés dans notre échantillon malgré leur vaccination : $p = \dfrac{43}{500} = 0,086$.
$p < 0,1$, donc **il semblerait** que le vaccin ait un effet protecteur. Pour autant, rien n'indique que nous ne soyons pas en présence d'un "échantillon exceptionnel". Un autre échantillon aurait pu donner des résultats très différents.

Comment donc justifier que ce que nous observons sur cet échantillon peut être extrapolé à toute la population ?

Vers une méthodologie

La première chose à faire est d'identifier ce que nous souhaitons tester, et comment le formuler mathématiquement.

Vers une méthodologie

La première chose à faire est d'identifier ce que nous souhaitons tester, et comment le formuler mathématiquement.

C'est bien évidemment l'efficacité du vaccin que le laboratoire souhaite mettre en évidence. On posera donc :

$$\begin{cases} H_0 & : \text{''Le vaccin n'est pas efficace''} \\ H_1 & : \text{''Le vaccin est efficace''} \end{cases}$$

Vers une méthodologie

La première chose à faire est d'identifier ce que nous souhaitons tester, et comment le formuler mathématiquement.

C'est bien évidemment l'efficacité du vaccin que le laboratoire souhaite mettre en évidence. On posera donc :

$$\begin{cases} H_0 & : \text{"Le vaccin n'est pas efficace"} \\ H_1 & : \text{"Le vaccin est efficace"} \end{cases}$$

Ce qui nous intéresse ici, c'est la proportion π de personnes de la population contractant la maladie si on vaccine toute la population.
Nous ne connaissons pas π car nous avons vacciné juste un échantillon.
Nous traduirons donc les hypothèses H_0 (hypothèse nulle) et H_1 (hypothèse alternative) par :

Vers une méthodologie

La première chose à faire est d'identifier ce que nous souhaitons tester, et comment le formuler mathématiquement.

C'est bien évidemment l'efficacité du vaccin que le laboratoire souhaite mettre en évidence. On posera donc :

$$\begin{cases} H_0 & : \text{"Le vaccin n'est pas efficace"} \\ H_1 & : \text{"Le vaccin est efficace"} \end{cases}$$

Ce qui nous intéresse ici, c'est la proportion π de personnes de la population contractant la maladie si on vaccine toute la population.
Nous ne connaissons pas π car nous avons vacciné juste un échantillon.
Nous traduirons donc les hypothèses H_0 (hypothèse nulle) et H_1 (hypothèse alternative) par :

$$\begin{cases} H_0 & : \pi = 0,10 \\ H_1 & : \pi < 0,10 \quad \text{(test unilatéral à gauche)} \end{cases}$$

Vers une méthodologie

Pour autant, une fois les hypothèses posées, quel processus va nous permettre de trancher entre H_0 et H_1 ?

Vers une méthodologie

Pour autant, une fois les hypothèses posées, quel processus va nous permettre de trancher entre H_0 et H_1 ?

L'erreur de première espèce α étant fixée, il faudra introduire une **variable de décision**, appelée aussi statistique du test, construite pour se servir des données de l'échantillon et répondre à la question posée.

Vers une méthodologie

Pour autant, une fois les hypothèses posées, quel processus va nous permettre de trancher entre H_0 et H_1 ?

L'erreur de première espèce α étant fixée, il faudra introduire une **variable de décision**, appelée aussi statistique du test, construite pour se servir des données de l'échantillon et répondre à la question posée.

Sa loi doit être parfaitement déterminée sous H_0 i.e en supposant H_0 vraie.

Méthode 1 avec les régions critiques

Méthode 1 : avec des valeurs critiques.

1. Nous énonçons deux hypothèses antagonistes H_0 et H_1 ;
2. On fixe a priori le risque $\alpha = P_{H_0 \text{ vraie}}(\text{Décider } H_1)$.
   Souvent $\alpha = 0;05$ ou $\alpha = 0,01$;
3. En supposant H_0 vraie, nous construisons une **variable de décision** D dite **statistique du test**.
4. Nous déterminons la ou les valeurs critiques d_c de rejet de H_0, ce qui nous donne les zones de rejet et de non rejet de H_0 ;
5. Nous calculons la valeur d_{obs} prise par la variable de décision D à l'aide des données de notre échantillon ;
6. Nous regardons ensuite où se situe cette valeur observée (dans la zone de rejet de H_0 ou pas) et nous concluons.

Méthode 2 avec les $p-$valeurs

Méthode 2 : utilisation des $p-$valeurs.

1. Nous énonçons deux hypothèses antagonistes H_0 et H_1 ;
2. On fixe a priori le risque $\alpha = P_{H_0 \text{ vraie}}(\text{Décider } H_1)$.
   Souvent $\alpha = 0;05$ ou $\alpha = 0,01$;
3. En supposant H_0 vraie, nous construisons une **variable de décision** D dite **statistique du test**.
4. Nous calculons la valeur d_{obs} prise par la variable de décision D à l'aide des données de notre échantillon ;
5. Nous calculons ensuite la probabilité p (sous H_0 toujours) que D prenne la valeur d_{obs} ou plus extrême.
6. Si $p < \alpha$, on rejette H_0, sinon on ne rejette pas H_0.

Avons-nous réellement extrapolé ?

Eh bien non ! Absolument pas !

Avons-nous réellement extrapolé ?

Eh bien non ! Absolument pas !

En effet, nous avons fait deux hypothèses H_0 et H_1 **sur un paramètre de la population**.

Et supposant H_0 vraie, nous avons évalué la vraissemblance que sous cette hypothèse, la variable de décision D prenne des valeurs plus extrêmes que d_{obs} (méthode des $p-$valeurs).

Avons-nous réellement extrapolé ?

Eh bien non ! Absolument pas !

En effet, nous avons fait deux hypothèses H_0 et H_1 **sur un paramètre de la population**.

Et supposant H_0 vraie, nous avons évalué la vraissemblance que sous cette hypothèse, la variable de décision D prenne des valeurs plus extrêmes que d_{obs} (méthode des $p-$valeurs).

Alors qu'extrapoler c'est exactement le contraire !
C'est évaluer la véracité de H_0 au vu des données de notre échantillon.
Bref, nous entrons dans le domaine des statistiques Bayésiennes (Hors programme).