

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

**UNIVERSITÉ MOHAMED KHIDER, BISKRA**

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

**DÉPARTEMENT DE Biologie**

Cours présenté par

Dr : Chine Amel

Module : Biostatistiques

Niveau : L2

2023-2024

# Chapitre 1

## Tests d'hypothèse

### 1.1 Généralités

**Définition 1** (*Tests statistique*) Les tests d'hypothèses (ou tests statistiques) sont un processus visant à fournir une règle de décision qui permet, sur la base des résultats d'un échantillon, de choisir entre deux hypothèses statistiques.

**Définition 2** (*Hypothèse nulle et hypothèse alternative*) L'hypothèse selon laquelle on fixe à priori un paramètre de la population à une valeur particulière s'appelle l'hypothèse nulle et est notée  $H_0$ .

l'hypothèse qui diffère de l'hypothèse  $H_0$  s'appelle l'hypothèse alternative (ou contre-hypothèse) et est notée  $H_1$ .

**Définition 3** (*Statistique*) Une statistique est une fonction des variables aléatoires représentant l'échantillon. Le choix de la statistique dépend de la nature des données, du type d'hypothèse que l'on désire contrôler.

**Définition 4** (*seuil de signification*) *Le seuil de signification  $\alpha$  est une probabilité comprise entre 0 et 1, donnée au début du test. Elle peut prendre les valeurs :0.05,0.01,0.1, ...*

**Définition 5** (*Région critique*) *La région critique, noté  $R_C$  correspond à l'ensemble des valeurs telles que :*

$$P(S \in R_C) = \alpha.$$

où  $S$  est la statistique du test.

**Définition 6** (*Règle de décision*) *la décision c'est à dire rejeter  $H_0$  ou l'accepter. Si on rejette  $H_0$  alors on accepte  $H_1$  et si on accepte  $H_0$  on rejette  $H_1$ . pour rejeter  $H_0$  ou l'accepter on compare la valeur de la statistique  $S$  avec  $S_{seuil}$  où  $P(S < S_{seuil}) = \alpha$ .*

## 1.2 Quelques tests

### 1.2.1 Test de conformité

Le test de conformité consiste à comparer entre un paramètre de la population avec un paramètre calculé sur l'échantillon, pour savoir si cet échantillon vient de cette population ou non. exemple de test de conformité, les tests sur la moyenne, la variance, ...

### 1.2.2 Test d'homogénéité

Le test d'homogénéité ou de comparaison consiste à vérifier que  $K(K \geq 2)$  échantillons (groupes) proviennent de la même population ou, cela revient à la même chose. C'est à dire on compare entre plus de deux échantillons pour savoir est ce que ces échantillons sont homogènes ou non.

### 1.2.3 Test d'indépendance

Le test d'indépendance consiste à éprouver l'existence d'une liaison entre deux variables. Les techniques utilisées diffèrent selon que les variables sont qualitatives nominales, ordinales ou quantitatives. Parmi les tests les plus utilisés lorsque la variable est qualitative c'est le test de Khi-deux.

### 1.2.4 Test de normalité

Le test de normalité consiste à vérifier la compatibilité des données avec la distribution loi normale, c'est à dire si l'échantillon suit la loi normale ou non.

## 1.3 Test de conformité de moyenne

Le test de conformité de moyenne consiste à comparer entre la moyenne de la population, noté  $\mu_0$  et la valeur empirique de la moyenne  $\bar{X}$ .

**Remarque 7** *Pour réaliser le test de conformité de moyenne il faut que la v.a étudiée suit la loi Normale  $N(\mu, \sigma^2)$*

*Il existe trois cas pour choisir la statistique de test :*

*pour  $n \geq 30$ , pour  $n < 30$  et  $\sigma^2$  connue et  $n < 30$  et  $\sigma^2$  inconnue*

*dans notre étude on prend seulement le troixième cas, c'est a dire pour  $n < 30$  et  $\sigma^2$  inconnue.*

**Exemple 8** *comme exemple, on veut savoir si la moyenne des notes de module biostat des étudiants de biologie égale à 12, ici  $\mu_0 = 12$*

**Exemple 9** *un autre exemple, on veut savoir si la moyenne de durée vie des lampes égale à 3 ans, ici  $\mu_0 = 3$ .*

### 1.3.1 Etapes pour réaliser le test de conformité de moyenne

On résume ce test dans les étapes suivantes :

**Etape 01 : les hypothèses** : les Hypothèses  $H_0$  et  $H_1$  de ce test sont définies par :

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

**Remarque 10** *Il existe d'autre écriture de l'hypothèse de ce test selon la formulation de la question, comme exemple*

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$

*ou*

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$$

*Mais dans notre exemples, on s'intresse toujours au premier cas, c'est à dire :*

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

**Etape 02 : la statistique** : La statistique, noté  $T$ , de ce test est suit la loi Student, elle

est définie par :

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \quad (1.1)$$

où :  $\bar{X}$  : est la moyenne empirique, définie par :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$$

et  $S$  : est l'écart type empirique, définie par :

$$S = \sqrt{S^2}$$

et  $S^2$  est la variance empirique, définie par :

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2.$$

**Etape 03 : la valeur critique :** Pour un seuil de signification  $\alpha$ , la valeur critique de ce test, noté  $T_\alpha$  est le fractile (quantile) d'ordre  $1 - \frac{\alpha}{2}$  et de degré de liberté  $n - 1$  :

$$T_\alpha = T_{n-1, 1-\frac{\alpha}{2}}$$

Cette valeur est lu de la table de Student.

**Etape 04 : la Décision :** Pour rejeter  $H_0$  ou l'accepter on compare la valeur  $T$  calculée avec la valeur théorique  $T_\alpha$  :

- Si  $T \in [-T_\alpha, T_\alpha]$  alors on accepte  $H_0$ , c'est à dire  $H_0$  est vraie.
- Si  $T \notin [-T_\alpha, T_\alpha]$  alors on rejette  $H_0$ , c'est à dire  $H_0$  est fausse.

**Exemple 11** *Un échantillon de 15 enfants d'une ville donnée à fourni les tailles suivantes (en cm) :*

70	85	93	99	101	105	110	121	138	166	74	85	93	99	102
----	----	----	----	-----	-----	-----	-----	-----	-----	----	----	----	----	-----

*On suppose que la v.a suit la loi Normale  $N(\mu, \sigma^2)$ .*

*Au vu de cet échantillon, peut-on considérer, au seuil de signification 2%, que la taille moyenne des enfants est de 110cm?*

On observe qu'on a un échantillon de taille  $n = 15 < 30$  et  $\sigma^2$  inconnue, la taille suit la loi Normale et on veut faire comparaison de moyenne où  $\mu_0 = 110$ , donc c'est le test de conformité de moyenne (test de Student).

**Etape 01 : les hypothèses** : les Hypothèses  $H_0$  et  $H_1$  de ce test sont définies par :

$$\begin{cases} H_0 : \mu = 110 \\ H_1 : \mu \neq 110 \end{cases}$$

**Etape 02 : la statistique** : on calcule la statistique par la formule (1.1)

1. On calcule :  $\bar{X}$

$$\bar{X} = \frac{1}{15} \sum_{i=1}^{15} x_i = 102.733$$

2. On calcule  $S^2$

$$S^2 = \frac{1}{14} \sum_{i=1}^{15} (x_i - 102.733)^2 = 598.924.$$

3. et  $S$  :

$$S = \sqrt{598.9234} = 24.473$$

4. et donc

$$T = \frac{102.733 - 110}{24.473/\sqrt{15}} = -1.1500$$

**Etape 03 : la valeur critique** : on a le seuil de signification  $\alpha = 0.02$ , alors la valeur critique, est le fractile (quantile) d'ordre  $1 - \frac{0.02}{2}$  et de degré de liberté  $n - 1$  :

$$T_\alpha = T_{n-1, 1-\frac{\alpha}{2}} = T_{14, 0.99} = 2.624$$

Cette valeur est lu de la table de Student. Voir (1.1)

**Etape 05 : la Décision** : On a  $-1.1500 \in [-2.624, 2.624]$  alors on accepte  $H_0$ , c'est à dire la taille moyenne des enfants est de 110cm.

## 1.4 Test d'homogénéité de moyennes pour $K=2$ (test de Student)

Le test d'homogénéité de moyennes consiste à comparer entre les moyennes des deux échantillons. .

### 1.4.1 Conditions nécessaires pour réaliser ce test

1. Les deux échantillons sont issus d'une loi normale
2. Les deux échantillons ont la même variance, c'est-à-dire il faut réaliser d'abord le test d'homogénéité de variances. ( $\sigma_1^2 = \sigma_2^2$ )
3. Les deux échantillons sont indépendants.

**Remarque 12** *Il existe trois cas pour choisir la statistique de test :*

*pour  $n_1 \geq 30, n_2 < 30$ , pour  $n_1 < 30, n_2 < 30$  et  $\sigma^2$  connue et  $n_1 < 30, n_2 < 30$  et  $\sigma^2$  inconnue*

*dans notre étude on prend seulement le troisième cas, c'est à dire pour  $n_1 < 30, n_2 < 30$  et  $\sigma^2$  inconnue.*

### 1.4.2 Etapes pour réaliser le test d'homogénéité de moyennes

On résume ce test dans les étapes suivantes :

**Etape 01 : les hypothèses** : les Hypothèses  $H_0$  et  $H_1$  de ce test sont définies par :

$$\left\{ \begin{array}{l} H_0 : \mu_1 = \mu_2 \text{ (les deux échantillons sont homogènes)} \\ H_1 : \mu_1 \neq \mu_2 \text{ (les deux échantillons sont non homogènes)} \end{array} \right.$$

**Etape 02 : la statistique** : La statistique, noté  $T$ , de ce test est suit la loi Student, elle

est définie par :

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (1.2)$$

où :  $\bar{X}$  : est la moyenne empirique, définie par :

$$\begin{aligned} \bar{X}_1 &= \frac{1}{n_1} \sum_{i=1}^{n_1} x_{1i} \\ \bar{X}_2 &= \frac{1}{n_2} \sum_{i=1}^{n_2} x_{2i} \end{aligned}$$

et  $S$  : est l'écart type empirique globale, définie par :

$$S = \sqrt{S^2}$$

et  $S^2$  est la variance empirique globale, définie par :

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \quad (1.3)$$

et

$$\begin{aligned} S_1^2 &= \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_{1i} - \bar{X}_1)^2 \\ S_2^2 &= \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (x_{2i} - \bar{X}_2)^2 \end{aligned}$$

**Etape 03 : la valeur critique :** Pour un seuil de signification  $\alpha$ , la valeur critique de ce test , noté  $T_\alpha$  est le fractile (quantile) d'ordre  $1 - \frac{\alpha}{2}$  et de degré de liberté  $n_1 + n_2 - 2$  :

$$T_\alpha = T_{n_1+n_2-2, 1-\frac{\alpha}{2}}$$

Cette valeur est lu de la table de Student.Voir (1.1)

**Etape 04 : la Décision :** Pour rejeter  $H_0$  ou l'accepter on compare la valeur  $T$  calculée avec la valeur théorique  $T_\alpha$  :

- Si  $T \in [-T_\alpha, T_\alpha]$  alors on accepte  $H_0$ , c'est à dire  $H_0$  est vraie.
- Si  $T \notin [-T_\alpha, T_\alpha]$  alors on rejette  $H_0$ , c'est à dire  $H_0$  est fausse.

**Exemple 13** Afin de comparer deux types d'arbre, nous avons réalisé un recueil de hauteur de quelques arbres, dont les mesures sont rangées dans le tableau suivant.

							Somme
Arbre 1	23.3	24.0	24.3	24.5	25.0	25.9	147
Arbre 2	21.1	21.1	22.1	22.4	23.3		110

- On suppose que la v.a suit la loi Normale  $N(\mu, \sigma^2)$ , les deux échantillons sont indépendants et les variances sont égaux.
- Que peut-on conclure sur la hauteur moyenne des deux types d'arbres à un seuil de risque  $\alpha = 2\%$  ?

On observe qu'on a deux échantillons de tailles  $n_1 = 6 < 30$  et  $n_2 = 5 < 30$  ainsi que  $\sigma^2$  est inconnue, donc c'est le test d'homogénéité de moyennes de Student

**Etape 01 : les hypothèses** : les Hypothèses  $H_0$  et  $H_1$  de ce test sont définies par :

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

**Etape 02 : la statistique** : on calcule la statistique par la formule (1.2)

1. On calcule :  $\bar{X}_1$  et  $\bar{X}_2$

$$\bar{X}_1 = \frac{1}{6} \sum_{i=1}^6 x_{1i} = \frac{147}{6} = 24.5$$

$$\bar{X}_2 = \frac{1}{5} \sum_{i=1}^5 x_{2i} = \frac{110}{5} = 22$$

2. On calcule  $S_1^2$  et  $S_2^2$  :

$$S_1^2 = \frac{1}{6} \sum_{i=1}^6 (x_{1i} - 24.5)^2 = 0.788.$$

$$S_2^2 = \frac{1}{5} \sum_{i=1}^5 (x_{2i} - 22)^2 = 0.870.$$

3. et  $S^2$  est la variance empirique globale par la formule 1.3

$$S^2 = \frac{(6-1)0.788 + (5-1)0.870}{6+5-2}$$

$$S^2 = 0.8244.$$

4. et

$$S = \sqrt{0.8244} = 0.908$$

5. et donc

$$T = \frac{24.5 - 22}{0.908(\frac{1}{6} + \frac{1}{5})} = 4.5469.$$

**Etape 03 : la valeur critique :** on a le seuil de signification  $\alpha = 0.02$ , alors la valeur critique, est le fractile (quantile) d'ordre  $1 - \frac{0.02}{2}$  et de degré de liberté  $6 + 5 - 2$  :

$$T_\alpha = T_{9, 1 - \frac{0.02}{2}} = T_{9, 0.99} = 2.821$$

Cette valeur est lu de la table de Student.

**Etape 05 : la Décision :** On a  $4.5469 \notin [-2.821, 2.821]$  alors on rejette  $H_0$ , c'est à dire les deux types d'arbres ont des hauteurs moyennes significativement différentes

LOI DE STUDENT AVEC  $k$  DEGRÉS DE LIBERTÉ  
QUANTILES D'ORDRE  $1 - \gamma$

$k$	$\gamma$										
	0.25	0.20	0.15	0.10	0.05	0.025	0.010	0.005	0.0025	0.0010	0.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.969
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
$\infty$	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

FIG. 1.1: Table de la loi de Student.