Comparaison de deux variances

Dr : RAIAH. M

Service de Biostatistique, Faculté de Médecine d'Oran

I. Introduction

On peut être amené à comparer deux variances dans les situations suivantes :

- Lorsque l'on veut vérifier la condition d'application d'égalité des variances dans le test T (test de Student) de comparaison de moyennes.
- Lorsque le problème porte sur la variabilité des observations.
  Exemple : choisir le meilleur instrument de mesure, celui qui a entre autre la variabilité la plus faible.

II. Test F de Fisher-Snedecor

- Sous l’hypothèse nulle H_0 les deux variances sont égales et leur rapport est égal à 1.
- Lorsque le rapport est significativement différent de 1, on rejette H_0 et on accepte l’hypothèse alternative H_1 qui suppose que les deux variances sont différentes.

III. Principe du test F de comparaison de deux variances

Le test F consiste :

- à calculer le rapport $F_O = S_A^2 / S_B^2$ (en mettant la plus grande variance au numérateur)
- Ce rapport est comparé à la valeur F lue sur la table F avec :
  - $ddl_A = k_A = (n_A - 1)$ et
  - $ddl_B = k_B = (n_B - 1)$ au point 2,5 %

- Si $\mathbf{F_0} \geq \mathbf{F_\alpha} \rightarrow H_0$ est rejetée , les variances sont significativement différentes.

- Si $\mathbf{F_0} < \mathbf{F_\alpha} \rightarrow H_0$ n'est pas rejetée, on peut supposer l'égalité des variances.

Condition d'application :

- les distributions doivent être supposées normales dans les deux populations d'où proviennent les deux échantillons.

Exemple :

On désire comparer la pression artérielle diastolique (PAD) d'un groupe de sujets sains (m_A = 70,1) et d'un groupe de sujets atteints de drépanocytose (m_B = 61,8).

On ne dispose que de 20 individus par groupe. La variance de la PAD est respectivement de 116,7 et de 57,2.

Solution :

Avant de mettre en place un test de Student, on vérifie ses conditions d'applications, en particulier l'égalité des variances.

Hypothèses :

- **H_0 :** les deux variances ne sont pas différentes.
- **H_1 :** les deux variances diffèrent.

$$F_o = 116,7/57,2 = 2,04$$

$k_A = 20 – 1 = 19$ et $k_B = 20 – 1 = 19$

On regarde dans la table $F_{2,5\%}$ à (19,19) d.d.l

La table de $F_{2,5\%}$ ne donne pas la valeur seuil pour des ddl égaux à 19, mais pour des ddl égaux à 20, $F_{2,5\%} = 2{,}46$.

La valeur trouvée F_o (2,04) est donc inférieure à la valeur seuil. On ne rejette pas H_0.

On peut donc raisonnablement comparer les deux moyennes de la PAD par un test T en supposant que la condition d'égalité des variances est remplie dans les deux populations.

α = 0,025

k_2 \ k_1	1	2	3	4	5	6	7	8	9	10	15	20	25	30	40	50	∞
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,43	6,33	6,27	6,23	6,18	6,14	6,02
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,27	5,17	5,11	5,07	5,01	4,98	4,85
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,57	4,47	4,40	4,36	4,31	4,28	4,14
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,10	4,00	3,94	3,89	3,84	3,81	3,67
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,77	3,67	3,60	3,56	3,51	3,47	3,33
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,52	3,42	3,35	3,31	3,26	3,22	3,08
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	2,86	2,76	2,69	2,64	2,59	2,55	2,40
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,57	2,46	2,40	2,35	2,29	2,25	2,09
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61	2,41	2,30	2,23	2,18	2,12	2,08	1,91
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,57	2,51	2,31	2,20	2,12	2,07	2,01	1,97	1,79
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39	2,18	2,07	1,99	1,94	1,88	1,83	1,64
50	5,34	3,97	3,39	3,05	2,83	2,67	2,55	2,46	2,38	2,32	2,11	1,99	1,92	1,87	1,80	1,75	1,55
100	5,18	3,83	3,25	2,92	2,70	2,54	2,42	2,32	2,24	2,18	1,97	1,85	1,77	1,71	1,64	1,59	1,35
∞	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,11	2,05	1,83	1,71	1,63	1,57	1,48	1,43	1,00