ISFA- M1
STATISTIQUE INFÉRENTIELLE

Fiche de TD N° 4 bis : COMPLÉMENTS ET INTERVALLES DE CONFIANCE.

1 Compléments.

Exercice 1.(Famille exponentielle)

Pour chacun des exercices de la feuille de TD N° 3, dire si le modèle statistique considéré fait partie de la famille exponentielle. Si c'est le cas, décrire les fonctions $\alpha(\theta), \beta(x), \gamma(\theta), \delta(x)$ et préciser une statistique exhaustive du modèle. Enfin, utiliser le théorème sur l'efficacité pour déterminer l'unique estimateur efficace (à transformation linéaire près) et son espérance $g(\theta)$.

Exercice 2.(Estimation en dimension 2)

Soit $X = (X_1, ..., X_n)$ un n-échantillon issu de la loi uniforme sur $[a, b]$ avec a et b, deux paramètres inconnus tels que $a < b$.

1. Déterminer l'estimateur du maximum de vraisemblance de (a, b). Est-ce une statistique exhaustive ? Peut-on appliquer le théorème sur les propriétés asymptotiques de l'EMV ?
2. Déterminer l'estimateur de (a, b) par la méthode des moments. Est-il convergent ?
3. * Montrer par la méthode Delta, qu'il est asymptotiquement normal et calculer sa variance asymptotique (Attention : longs calculs ! On pourra se contenter d'exprimer la matrice de covariance de (X_1, X_1^2) ainsi que la matrice jacobienne de la fonction g de la méthode Delta.)

2 Intervalles de confiance (suite)

Exercice 3. (Intervalle de confiance asymptotique : suite de l'exercice 3 de la feuille 4)

On reprend les notations de l'exercice 3 de la feuille de TD N° 4. Sur un échantillon représentatif de 1000 personnes, on étudie l'évolution des avis favorables pour le Maire. En novembre, il y avait $F_n = 43\%$ d'avis favorables et en décembre, la côte baisse à $F'_n = 41\%$. Le Maire doit-il prendre cette baisse de popularité au sérieux ?

1. On suppose observer deux n-échantillons indépendants $X = (X_1, ..., X_n)$ et $X' = (X'_1, ..., X'_n)$ issus de lois de Bernoulli de paramètre respectifs p et p'. Montrer que l'estimateur $F_n - F'_n$ est asymptotiquement normal.
2. Déterminer un intervalle de confiance de niveau 95% de type $[a_n, +\infty[$ pour $p - p'$. Faire le calcul pour $n = 1000$.
3. Peut-on conclure à la baisse de popularité du Maire (i.e $p' < p$) ?

* 4) Reformuler la question précédente sous forme de test ?

<u>Exercice 4</u>.**(Intervalles de confiance pour la loi normale)**

On s'interroge sur la comparaison des tailles moyennes des garçons et des filles de 6 ans dans une population, pour cela on a pris comme échantillon, jugé représentatif de cette tranche d'âge, une classe d'école primaire (niveau CP en France), et on a observé :

16 garçons : moyenne 126.5 cm, écart-type estimé 12.9 cm

15 filles : moyenne 136.9 cm, écart-type estimé 11.9 cm.

On admet que la distribution des tailles dans chacune des sous-populations (garçons, filles) suit une loi gaussienne.

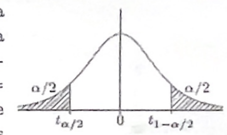
1. Donner des intervalles de confiance à 95% pour les tailles moyennes des garçons et des filles.
2. Donner un intervalle de confiance à 95% pour l'écart-type de la taille des garçons. Même question pour les filles.

3) Donner un intervalle de confiance à 95% pour le rapport des deux écart-types. Peut-on conclure que les variances des deux populations sont différentes ? (On donne, pour la loi de Fisher-Snedecor à (15,14) degrés de libertés, les quantiles $f_{0.975} = 2.95$ et $f_{0.025} = \frac{1}{f_{0.975}} = 0.339$).

* 4) Sur la base de la réponse à la question précédente, on suppose que la variance est la même dans les deux populations. Par ailleurs, au vu de cet échantillon, un observateur avance l'opinion : dans la population, la taille moyenne des filles dépasse de plus de 2 cm celle des garçons. Les données confirment-elles significativement, au niveau $\alpha = 0.05$, cette opinion ? (autrement dit quelle est la conclusion, au niveau $\alpha = 0.05$, du test de l'hypothèse nulle : dans la population, la taille moyenne des filles dépasse de moins de 2 cm celle des garçons ?).

Quantiles de la loi de Student.

Si T est une variable aléatoire suivant la loi de Student à ν degrés de liberté, la table suivante donne, pour α fixé, la valeur $t_{1-\alpha/2}$ telle que $\mathbb{P}(|T| \geq t_{1-\alpha/2}) = \alpha$. Ainsi, $t_{1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ de la loi de Student à ν degrés de liberté.



ν \ α	0,900	0,500	0,300	0,200	0,100	0,050	0,020	0,010	0,001
1	0,1584	1,0000	1.9626	3,0777	6.3138	12,7062	31,8205	63,6567	636.6193
2	0,1421	0,8165	1.3862	1,8856	2.9200	4,3027	6,9646	9,9248	31,5991
3	0.1366	0,7649	1,2498	1,6377	2,3534	3,1824	4,5407	5.8409	12,9240
4	0.1338	0,7407	1.1896	1,5332	2,1318	2,7764	3,7469	4,6041	8,6103
5	0,1322	0,7267	1,1558	1,4759	2,0150	2,5706	3.3649	4,0321	6,8688
6	0,1311	0,7176	1.1342	1,4398	1,9432	2,4469	3,1427	3,7074	5,9588
7	0,1303	0.7111	1.1192	1,4149	1.8946	2,3646	2,9980	3,4995	5,4079
8	0,1297	0,7064	1,1081	1,3968	1,8595	2,3060	2,8965	3,3554	5,0413
9	0,1293	0,7027	1,0997	1,3830	1,8331	2,2622	2,8214	3,2498	4,7809
10	0,1289	0,6998	1,0931	1,3722	1,8125	2,2281	2,7638	3,1693	4,5869
11	0,1286	0,6974	1,0877	1,3634	1,7959	2,2010	2,7181	3,1058	4,4370
12	0,1283	0,6955	1,0832	1,3562	1,7823	2,1788	2,6810	3,0545	4,3178
13	0,1281	0,6938	1,0795	1,3502	1,7709	2,1604	2,6503	3,0123	4,2208
14	0,1280	0,6924	1,0763	1,3450	1,7613	2,1448	2,6245	2,9768	4,1405
15	0,1278	0.6912	1.0735	1,3406	1,7531	2,1314	2,6025	2,9467	4,0728
16	0,1277	0,6901	1,0711	1,3368	1,7459	2,1199	2,5835	2,9208	4,0150
17	0,1276	0,6892	1.0690	1,3334	1,7396	2,1098	2,5669	2,8982	3,9651
18	0,1274	0,6884	1.0672	1,3304	1.7341	2,1009	2,5524	2,8784	3,9216
19	0,1274	0,6876	1.0655	1,3277	1,7291	2,0930	2,5395	2.8609	3,8834
20	0,1273	0,6870	1,0640	1,3253	1,7247	2,0860	2,5280	2,8453	3,8495
21	0,1272	0,6864	1,0627	1,3232	1,7207	2,0796	2,5176	2,8314	3,8193
22	0,1271	0,6858	1,0614	1,3212	1,7171	2,0739	2.5083	2,8188	3,7921
23	0,1271	0,6853	1,0603	1,3195	1,7139	2,0687	2,4999	2,8073	3,7676
24	0,1270	0,6848	1,0593	1,3178	1,7109	2,0639	2,4922	2,7969	3,7454
25	0,1269	0,6844	1.0584	1,3163	1,7081	2,0595	2,4851	2,7874	3,7251
26	0,1269	0,6840	1.0575	1,3150	1.7056	2,0555	2,4786	2,7787	3,7066
27	0,1268	0,6837	1,0567	1,3137	1,7033	2,0518	2,4727	2,7707	3,6896
28	0,1268	0,6834	1,0560	1,3125	1,7011	2,0484	2,4671	2,7633	3,6739
29	0,1268	0,6830	1,0553	1,3114	1,6991	2,0452	2,4620	2,7564	3,6594
30	0,1267	0,6828	1,0547	1,3104	1,6973	2,0423	2,4573	2,7500	3,6460
40	0,1265	0,6807	1,0500	1,3031	1,6839	2,0211	2,4233	2,7045	3,5510
60	0,1262	0,6786	1,0455	1,2958	1,6706	2,0003	2,3901	2,6603	3,4602
80	0,1261	0,6776	1.0432	1,2922	1.6641	1,9901	2,3739	2,6387	3,4163
120	0,1259	0,6765	1.0409	1,2886	1,6577	1,9799	2,3578	2,6174	3,3735
∞	0,1257	0,6745	1.0364	1,2816	1,6449	1,9600	2,3263	2,5758	3,2905

Exercice 2 :

$X_i \sim \mathcal{U}[a,b]$

1) $L(X_1, \dots, X_n, a, b) = \prod \frac{1}{b-a} \mathbb{1}_{X_i \geq a} \mathbb{1}_{X_i \leq b} = (b-a)^{-n} \mathbb{1}_{\min X_i \geq a} \mathbb{1}_{\max X_i \leq b}$

non indep du support $\Rightarrow \begin{cases} \hat{a}_n = \min X_i \\ \hat{b}_n = \max X_i \end{cases}$

Oui stat exhaustive car $L(X_1, \dots, X_n, a, b) = (b-a)^{-n} \mathbb{1}_{\hat{a}_n \geq a} \mathbb{1}_{\hat{b}_n \leq b}$

Non car il n'y a pas indep du support $\Rightarrow$ On ne peut pas dire

$\sqrt{n}\begin{pmatrix} \hat{a}_n - a \\ \hat{b}_n - b \end{pmatrix} \longrightarrow \mathcal{N}(0, I^{-1})$

2) $\mathbb{E}[X_i] = \frac{a+b}{2}$

$\hat{a}_M =$

$V(X_i) = \frac{(b-a)^2}{12}$

$\Rightarrow \begin{cases} \frac{\hat{a}_M + \hat{b}_M}{2} = \overline{X}_n \\ \frac{(\hat{b}_M - \hat{a}_M)^2}{12} = \overline{X_n^2} - \overline{X}_n^2 \end{cases} \Rightarrow \begin{cases} \hat{a}_M = 2\overline{X}_n - \hat{b}_M \\ \frac{(2\hat{b}_M - 2\overline{X}_n)^2}{12} = \overline{X_n^2} - \overline{X}_n^2 \end{cases}$

$\Rightarrow \begin{cases} \hat{a}_M = \overline{X}_n - \sqrt{3}\sqrt{\overline{X_n^2} - \overline{X}_n^2} \\ \hat{b}_M = \overline{X}_n + \sqrt{3}\sqrt{\overline{X_n^2} - \overline{X}_n^2} \end{cases}$