

EXAMEN D'ANALYSE STATISTIQUE

Deuxième Session

Question de cours :

Soit un échantillon simple au hasard  $(X_1, X_2, \dots, X_n)$ . Comment dénomme-t-on l'expression définie par  $\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  ? Est-ce un estimateur sans biais de quelque chose ? (*pas de démonstration*)

Exercice 1 : Fonctions de densité, répartition et quantiles

Soit la fonction de densité pour une variable aléatoire X continue définie sur  $[0 ; 30]$  par

$$f(x) = \begin{cases} \frac{k}{2} & \text{si } x \leq 20 \\ k & \text{si } x > 20 \end{cases}$$

- 1) Tracez cette fonction puis à l'aide d'un raisonnement géométrique indiquez quelle valeur doit prendre k pour que la fonction f soit une fonction de densité.
- 2) Donnez l'expression de F, la fonction de répartition de f.
- 3) Tracez la fonction de répartition
- 4) La loi est-elle symétrique ? Quelle est la médiane ?
- 5) Indiquez la valeur des quantiles Q20% et Q80%
- 6) Calculez  $Q5\% = F^{-1}(0,05)$  et interprétez.

Exercice 2 : Sondages

Afin de prévoir les résultats d'une élection avec 2 candidats A et B,, on réalise un sondage d'opinion dans une ville constituée de 4100 électeurs. Les préférences électorales issues d'un échantillon simple au hasard se répartissent ainsi :

$X_i$	Candidat A ( $X = 0$ )	Candidat B ( $X = 1$ )
$Y_j$		
Cadres ( $Y = y_1$ )	40	28
Professions intermédiaires ( $Y=y_2$ )	142	46
Employés – Ouvriers ( $Y = y_3$ )	352	344

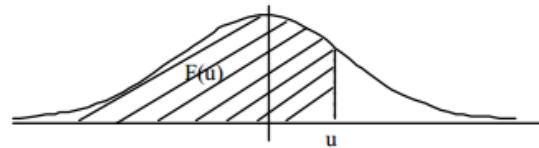
## Partie I : Intervalle de confiance des intentions de vote en faveur du candidat B

- 1) En utilisant un théorème vu en cours, indiquez quelle est la loi de probabilité approchée de la Fréquence Empirique lorsque la taille de l'échantillon tend vers l'infini ? (*preuve*)
- 2) A partir de la réponse à la question précédente, démontrez la formule de l'intervalle de confiance de niveau  $1-\alpha$  dans le cas d'une variable aléatoire parente X discrète prenant les valeurs 0 ou 1.
- 3) Appliquez cette formule en donnant un intervalle de confiance des intentions de vote en faveur du candidat B de niveau 90%.
- 4) Interprétez. En choisissant un seuil d'erreur de 10%, répondez à la question suivante : le candidat B va-t-il remporter les élections ?

## Partie II: Dépendance entre profession et intentions de vote

- 5) En utilisant les estimateurs des probabilités conditionnelles, indiquez si les cadres soutiennent moins le candidat B que les professions intermédiaires (comparaison directe sans test)
- 6) Réalisez à présent un test du Khi-deux indiquant s'il y a indépendance entre la catégorie socio-professionnelle (CSP) et les préférences électorales.

### Variable NORMALE CENTREE REDUITE



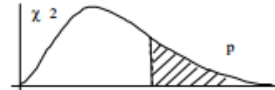
$$U \approx N(0,1) \quad P(U \leq u) = F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-x^2/2} dx$$

$$F(-u) = 1 - F(u) \quad P(|U| \leq u) = 2 F(u) - 1$$

TABLE de F(u) en fonction de u :

u	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767

TABLE DU CHI-DEUX :  $\chi^2(n)$



n P	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.02	0.01
1	0,0158	0,0642	0,148	0,455	1,074	1,642	2,706	3,841	5,412	6,635
2	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	7,824	9,210
3	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	9,837	11,341
4	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	11,668	13,277
5	1,610	2,343	3,000	4,351	6,064	7,289	9,236	11,070	13,388	15,086
6	2,204	3,070	3,828	5,348	7,231	8,558	10,645	12,592	15,033	16,812
7	2,833	3,822	4,671	6,346	8,383	9,803	12,017	14,067	16,622	18,475
8	3,490	4,594	5,527	7,344	9,524	11,030	13,362	15,507	18,168	20,090
9	4,168	5,380	6,393	8,343	10,656	12,242	14,684	16,919	19,679	21,666
10	4,865	6,179	7,267	9,342	11,781	13,442	15,987	18,307	21,161	23,209
11	5,578	6,989	8,148	10,341	12,899	14,631	17,275	19,675	22,618	24,725
12	6,304	7,807	9,034	11,340	14,011	15,812	18,549	21,026	24,054	26,217
13	7,042	8,634	9,926	12,340	15,119	16,985	19,812	22,362	25,472	27,688
14	7,790	9,467	10,821	13,339	16,222	18,151	21,064	23,685	26,873	29,141
15	8,547	10,307	11,721	14,339	17,322	19,311	22,307	24,996	28,259	30,578
16	9,312	11,152	12,624	15,338	18,418	20,465	23,542	26,296	29,633	32,000
17	10,085	12,002	13,531	16,338	19,511	21,615	24,769	27,587	30,995	33,409
18	10,865	12,857	14,440	17,338	20,601	22,760	25,989	28,869	32,346	34,805
19	11,651	13,716	15,352	18,338	21,689	23,900	27,204	30,144	33,687	36,191
20	12,443	14,578	16,266	19,337	22,775	25,038	28,412	31,410	35,020	37,566
21	13,240	15,445	17,182	20,337	23,858	26,171	29,615	32,671	36,343	38,932
22	14,041	16,314	18,101	21,337	24,939	27,301	30,813	33,924	37,659	40,289
23	14,848	17,187	19,021	22,337	26,018	28,429	32,007	35,172	38,968	41,638
24	15,659	18,062	19,943	23,337	27,096	29,553	33,196	36,415	40,270	42,980
25	16,473	18,940	20,867	24,337	28,172	30,675	34,382	37,652	41,566	44,314
26	17,292	19,820	21,792	25,336	29,246	31,795	35,563	38,885	42,856	45,642
27	18,114	20,703	22,719	26,336	30,319	32,912	36,741	40,113	44,140	46,963
28	18,939	21,588	23,647	27,336	31,391	34,027	37,916	41,337	45,419	48,278
29	19,768	22,475	24,577	28,336	32,461	35,139	39,087	42,557	46,693	49,588
30	20,599	23,364	25,508	29,336	33,530	36,250	40,256	43,773	47,962	50,892

Pour  $n > 30$ , on peut admettre que  $\sqrt{2\chi^2} - \sqrt{2n-1} \approx N(0,1)$