Épreuve de Statistique IV

Problème (17pts): dans une certaine région montagneuse, un représentant de la Fédération de la Randonnée Pédestre éditant les Topo-Guides TG2 s'est livré à une petite enquête pour connaître le temps de montée d'un randonneur. Les résultats obtenus sont dans le tableau suivant:

Temps de montée en minutes	[50-60[	[60-70[	[70-90[	[90-100[	[100-120]
Effectif	2	8	25	10	5

1) (5pts): En supposant la normalité du temps de montée, tester l'hypothèse que le temps moyen de montée est de 80 mn au risque de 5%.

2) (3pts): En supposant toujours la normalité du temps de montée, tester, avec un risque de 5%, l'hypothèse H_0 : " $\sigma^2 = 225$ " contre l'hypothèse H_1 : " $\sigma^2 > 225$ ".

3) (5pts): Tester cette hypothèse de normalité faites en 1) et 2), au risque de 5% (moyenne-variance inconnues).

4) (4pts): En effectuant son enquête il a aussi demandé aux randonneurs si le TG2 était leur chemin préféré pour monter au sommet de la montagne. 32 des personnes interrogées lui ont répondu :**oui**.
Tester l'hypothèse que plus de la moitié des randonneurs préfère emprunter le TG2 avec un risque de 5%.

Calculer la *p-value* et conclure.

Exercice (3pts): Un système assez sommaire de remplissage de doses a été fabriqué. Pour limiter les risques, 3 doses sont fabriquées chaque matin entre 7 et 8 h. N'ayant pas d'instrument suffisamment précis pour mesurer le volume, le médecin pèse chaque dose fabriquée, et note la moyenne et l'étendue des poids des 3 doses. Au bout de 30 jours de fabrication, la moyenne des 30 moyennes est égale à 3,45 et la moyenne des 30 étendues à 0,32.
1) Déterminer les limites de contrôle des cartes de contrôle de la moyenne et de l'étendue.
2) Quel est le nombre moyen d'échantillons à prélever avant de détecter un dérèglement du processus de remplissage conduisant à une déviation de la moyenne de 1,5écarts-type?

N.B

- ***arrondir tous les résultats à la*** deuxième ***décimale !***
- ***Vous aurez besoin de certaines données parmi les suivantes :***

➤ Si ***X(n)*** est une variable aléatoire de ***Khi-deux*** à ***n*** degrés de liberté, on a :

$P(X(49) > 66{,}34) = 0{,}05$ $P(X(49) > 70{,}22) = 0{,}025$ et $P(X(50) > 71{,}42) = 0{,}025$.

$P(X(1) > 3{,}84) = 0{,}05$ $P(X(2) > 5{,}99) = 0{,}05$ $P(X(3) > 7{,}81) = 0{,}05$ et $P(X(4) > 9{,}49) = 0{,}05$.

➤ Si ***T(n)*** est une variable aléatoire de ***Student*** à ***n*** degrés de liberté, on a :

P(T(49) >2) =0,025 et P(T(50) >1,68) =0,05 et   P(T(48) > 2,01) = 0,025.

➢ Si ***Z*** suit la *loi normale centrée réduite* : u sont les valeurs critiques de Z **:** ***P(Z > u) = α et P(Z > 2,6) = 0 .***

u	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07
0	0,50	0,50	0,49	0,49	0,48	0,48	0,48	0,47
0,1	0,46	0,46	0,45	0,45	0,44	0,44	0,44	0,43
0,2	0,42	0,42	0,41	0,41	0,41	0,40	0,40	0,39
0,3	0,38	0,38	0,37	0,37	0,37	0,36	0,36	0,36
0,4	0,34	0,34	0,34	0,33	0,33	0,33	0,32	0,32
0,5	0,31	0,31	0,30	0,30	0,29	0,29	0,29	0,28
0,6	0,27	0,27	0,27	0,26	0,26	0,26	0,25	0,25
0,7	0,24	0,24	0,24	0,23	0,23	0,23	0,22	0,22
0,8	0,21	0,21	0,21	0,20	0,20	0,20	0,19	0,19
0,9	0,18	0,18	0,18	0,18	0,17	0,17	0,17	0,17
1	0,16	0,16	0,15	0,15	0,15	0,15	0,14	0,14
1,1	0,14	0,13	0,13	0,13	0,13	0,13	0,12	0,12
1,2	0,12	0,11	0,11	0,11	0,11	0,11	0,10	0,10
1,3	0,10	0,10	0,09	0,09	0,09	0,09	0,09	0,09
1,4	0,08	0,08	0,08	0,08	0,07	0,07	0,07	0,07
1,5	0,07	0,07	0,06	0,06	0,06	0,06	0,06	0,06
1,6	0,06	0,06	0,06	0,06	0,06	0,05	0,04	0,04
1,7	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,04
1,8	0,04	0,04	0,03	0,03	0,03	0,03	0,03	0,03
1,9	0,03	0,03	0,03	0,03	0,03	0,03	0,02	0,02
2	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
2,1	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
2,2	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
2,3	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
2,4	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
2,5	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01

TABLEAU DES COEFFICIENTS POUR LE CALCUL DES CARTES

n	2	3	4	5	6	7	8	9	10
d_2	1,13	1,69	2,06	2,33	2,53	2,7	2,85	2,97	3,08
A_2	1,88	1,02	0,73	0,58	0,48	0,42	0,37	0,34	0,31
D_3	0	0	0	0	0	0,07	0,14	0,18	0,22
D_4	3,27	2,57	2,28	2,11	2	1,92	1,86	1,82	1,78

Université Sidi Mohamed Ben Abdellah
Faculté des sciences Juridiques,
Economiques et Sociales
- Fès-

Prof : TOUIJAR

Année Universitaire 2007/2008
Filière: Sciences Economiques et Gestion
Module: Méthodes quantitatives IV
S5 - Sections **C & D**

Eléments de correction de l'épreuve de Statistique IV

Problème (17pts):

1- a) F.H. $\begin{cases} H_0: & \mu = \mu_0 (= 80) \\ \# & \\ H_1: & \mu \neq \mu_0 \end{cases}$ test de moyenne du genre TB

b) paramètre : μ

c) Statistique :$\bar{X}$ « bon» estimateur de μ

d) loi de $\bar{X}$ sous H_0 :

Hyp de Normalité et σ^2 inconnue $\Rightarrow$ $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim \boldsymbol{t}(n-1)$

e) R.D.

$$\begin{cases} si \quad |\bar{x} - \mu_0| > c_{\alpha/2} = \boldsymbol{t}_{\alpha/2}(n-1)\frac{s}{\sqrt{n}} & on\ rejette\ H_0 \\ si \quad |\bar{x} - \mu_0| \leq c_{\alpha/2} = \boldsymbol{t}_{\alpha/2}(n-1)\frac{s}{\sqrt{n}} & on\ ne\ peut\ rejeter\ H_0 \end{cases}$$

f) A.N.

Temps de montée	n_i	x_i	$n_i x_i$	$n_i x_i^2$
[50-60[	2	55	110	6050
[60-70[	8	65	520	33800
[70-90[	25	80	2000	160000
[90-100[	10	95	950	90250
[100-120]	5	110	550	60500
TOTAL	50	-	4130	350600

$$\bar{x} = \frac{1}{50}\sum_{i=1}^{5} n_i x_i = \frac{4130}{50} = 82{,}6\ mn \qquad t_{0,025}(49) = 2$$

$$s^2 = \frac{n}{n-1} s_e^2 = \frac{50}{49}\left(\overline{x^2} - \bar{x}^2\right) = \frac{50}{49}\left(\frac{350600}{50} - 82{,}6^2\right) = 193{,}1$$

$$\Rightarrow s = 13{,}9\ mn \Longrightarrow c_{\alpha/2} = 3{,}93 \quad et \quad |\bar{x} - \mu_0| = |82{,}6 - 80| = 2{,}6$$

$$d'où \quad |\bar{x} - \mu_0| < c_{\alpha/2} \Longrightarrow \mathrm{NRH}_0$$

Conclusion : les données ne nous permettent pas de rejeter l'hypothèse que le temps moyen de la montée est de 80 mn.

2- a) paramètre : σ^2

b) F.H. $\begin{cases} H_0: & \sigma^2 = \sigma_0^2 = 250 \\ \# & \\ H_1: & \sigma^2 > \sigma_0^2 \end{cases}$ test de variance du genre TUD

c) Statistique : S^2 « bon» estimateur de σ^2

d) loi de S^2 sous H_0 :

Hyp de Normalité $\Rightarrow$ $Y = \frac{(n-1)S^2}{\sigma_0^2} \rightsquigarrow \chi 2(n-1)$

e) R.D.

$$\begin{cases} si \quad s^2 > c_\alpha = \frac{\sigma_0^2}{(n-1)} \chi_\alpha^2(n-1) & on\ rejette\ H_0 \\ si \quad s^2 \le c_\alpha = \frac{\sigma_0^2}{(n-1)} \chi_\alpha^2(n-1) & on\ ne\ peut\ rejeter\ H_0 \end{cases}$$

f) A.N.

$s^2 = 193{,}1$ (déjà vu en 1)) et $\chi^2_{0,05}(49) = 66{,}34$

$$donc\ c_\alpha = \frac{225}{49} \times 66{,}34 = 304{,}62$$

$$d'où \quad s^2 \le c_\alpha \Rightarrow NRH_0$$

Conclusion : les données ne nous permettent pas de rejeter l'hypothèse que la variance du temps de la montée est de 225.

3- a) F.H. $\begin{cases} H_0: & X \sim \mathcal{N}(\mu, \sigma^2) \\ \# & \\ H_1: & X \not\sim \mathcal{N}(\mu, \sigma^2) \end{cases}$ test d'ajustement de χ^2 à la loi normale

b) Statistique : $D^2 = \sum_{i=1}^{k} \frac{(N_i - n_{ti})^2}{n_{ti}}$

c) Loi de D^2 sous H_0 : $D^2 \sim \chi^2_{k-m-1}$; où m est le nombre de paramètres à estimer.

d) R.D.

$$\begin{cases} Si\ \ d^2 > \chi^2_{k-m-1;\alpha} & on\ rejette\ H_0 \\ Si\ \ d^2 \le \chi^2_{k-m-1;\alpha} & on\ ne\ rejette\ pas\ H_0 \end{cases}$$

e) A.N. : on a deux paramètres à estimer μ (moyenne) et σ^2 (variance) ; ce qui est déjà vu en 1)

- Estimation de μ et de σ^2 :

$$\hat{\mu} = \bar{x} = 82{,}6\ mn \quad \text{et} \quad \hat{\sigma}^2 = s^2 = 193{,}1 \Rightarrow \hat{\sigma} = 13{,}9\ mn$$

Calculons les effectifs théoriques $n_{ti} = np_i$

- Tableau théorique : $Z_i = \frac{e_i - \hat{\mu}}{\hat{\sigma}} \sim \mathcal{N}(0,1)$

$[e_{i-1} ; e_i[$	$[z_{i-1} ; z_i[$	p_i	n_{ti}	n_i	d^2
[50 ; 60[	]-∞ ; -1,63[	0,06	3 } 9	10	0,11
[60 ; 70[	[-1,63 ; -0,91[	0,12	6		
[70 ; 90[	[-0,91 ; 0,53[	0,52	26	25	0,04
[90 ; 100[	[0,53 ; 1,25[	0,19	9,5	10	0,03
[100 ; 120]	[1,25 ; +∞[	0,11	9,5	5	0,05
TOTAL	**TOTAL**	**1,00**	**50**	**50**	**0,23**

On a $d^2 = 0{,}23$ et $\chi^2_\alpha(4-2-1) = \chi^2_\alpha(1) = 3{,}84$

D'où $d^2 < \chi^2_\alpha(1) \Rightarrow \text{NRH}_0$

f)- **Conclusion :** les données ne nous permettent pas de rejeter l'hypothèse que les données proviennent d'une population Normale.

4- a) paramètre : p

b) F.H. $\begin{cases} H_0 : \ p = p_0 = 0{,}5 \\ \# \\ H_1 : \ p > p_0 = 0{,}5 \end{cases}$ test de proportion du genre TUD

c) Statistique : F « bon» estimateur de p

d) loi de F sous H_0 :

T.C.L.$\Rightarrow n = 50 \geq 30 \quad np_0 = nq_0 = 25 \geq 5$

d'où $F \approx \mathcal{N}\left(p_0 ; \frac{p_0 q_0}{n}\right) \quad \Rightarrow \quad Z = \frac{F - p_0}{\sqrt{p_0 q_0 / n}} \approx \mathcal{N}(0;1)$

e) R.D.

$$\begin{cases} si \quad f > c_\alpha = p_0 + z_\alpha \sqrt{p_0 q_0 / n} \quad on\ rejette\ H_0 \\ si \quad f \leq c_\alpha = p_0 + z_\alpha \sqrt{p_0 q_0 / n} \quad on\ ne\ peut\ rejeter\ H_0 \end{cases}$$

f)A.N. $f = \frac{32}{50} = 0{,}64 \quad z_{0,05} = 1{,}65 \quad c_\alpha = 0{,}5 + 1{,}65\sqrt{\frac{0{,}25}{50}} = 0{,}62$

$d'où \quad f > c_\alpha \Longrightarrow RH_0$

Conclusion : les données semblent être cohérentes avec l'hypothèse que plus de la moitié des randonneurs préfère emprunter le TG2.

$$\alpha_0 = \mathbb{p}_0(F > f) = \mathbb{p}_0\left(Z > \frac{f - p_0}{\sqrt{p_0 q_0 / n}}\right) = \mathbb{p}_0(Z > 2) = 2\%$$

Le test est donc assez significatif.

Exercice (3pts): 1)

- *μ et σ inconnues*
- ***a)*** *les limites de contrôle sur la carte* $\overline{X}$: $n=3$; $\bar{\bar{x}} = 3{,}45$; $\bar{r} = 0{,}32$
- $UCL = \bar{\bar{x}} + 3\widehat{\sigma_{\bar{x}}} = \bar{\bar{x}} + A_2\bar{r} = 3{,}45 + 1{,}02 * 0{,}32 = 3{,}78$
- $LCL = \bar{\bar{x}} - 3\widehat{\sigma_{\bar{x}}} = \bar{\bar{x}} - A_2\bar{r} = 3{,}45 - 1{,}02 * 0{,}32 = 3{,}12$
- CL=3,45
- ***b)*** *les limites de contrôle sur la carte des étendues:*
- $UCL = D_4\bar{r} = 2{,}57 * 0{,}32 = 0{,}82$
- $LCL = D_3\bar{r} = 0 * 0{,}32 = 0$
- CL=0,32

2) $\beta = P_1\left(-3 - 1{,}5\sqrt{3} < Z_1 < +3 - 1{,}5\sqrt{3}\right) = P_1\left(Z_1 < +0{,}4\right) = 0{,}66$

$$\Rightarrow POM_{1,5;3} = \frac{1}{1-\beta} \cong 3$$

N.B. : **On n'accepte pas les formules sans application numérique (i.e. sans faire de calcul !)**