

Examen de première session. Durée : 2h

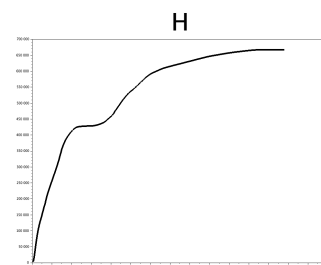
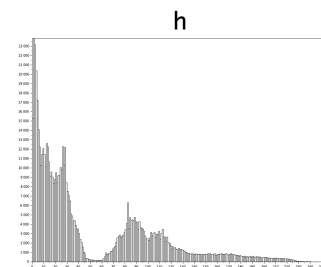
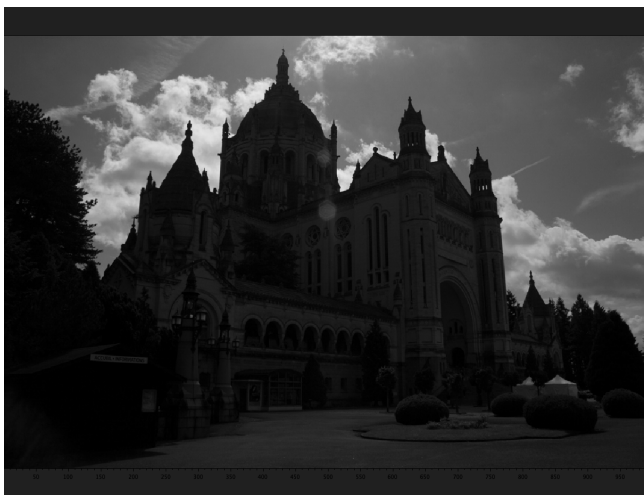
Les différents exercices sont indépendants et peuvent être traités dans un ordre quelconque.

La note de l'examen sera la moyenne sur 20 des deux thèmes.

Thème image : 20 points

Exercice 1 : Histogrammes et correction d'image (10 points)

1. Rappeler la définition de l'histogramme d'une image.
2. Rappeler la définition de l'histogramme cumulé d'une image.
3. Rappeler le principe de l'égalisation d'histogramme et son effet sur ces deux histogrammes.
4. On considère l'image ci-dessous, ainsi que son histogramme h et son histogramme cumulé H .



Commenter les allures de h et H en lien avec l'image.

Afin d'améliorer son contraste, on lui applique deux types de transformation : la transformation affine $T1(p) = 2p$ et la transformation $T2(p) = 255 H(p)/H(255)$ qui correspond à l'égalisation d'histogramme.

5. Examinez les images A et B ci-dessous et indiquez laquelle correspond à la transformation $T1$ et à la transformation $T2$. Justifiez votre choix.

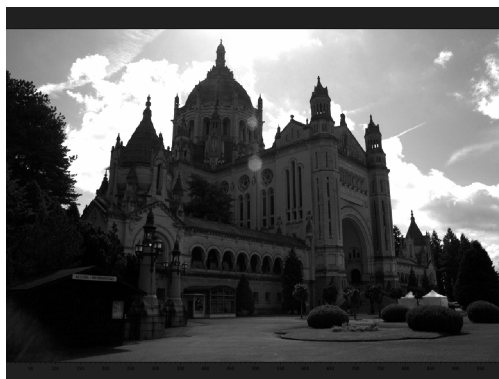


Image A



Image B

6. Donnez l'allure des histogrammes des images A et B en justifiant votre dessin. Même question pour les histogrammes cumulés.

Exercice 2 : Réhaussement des contours (10 points)

On considère le filtre $H = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$, dont l'effet sur une image est d'augmenter sa netteté apparente :



Image initiale



Image transformée

Le but de cet exercice est de comprendre son fonctionnement à partir d'un exemple simple.

Soit $I1$ une image de taille $M \times N$ et soit $I2 = \text{conv}(I1, H)$ l'image produite par convolution de $I1$ avec H ; on la définit en `scilab` de la manière suivante : pour tout u allant de 2 à $M-1$, et pour tout v allant de 2 à $N-1$,

$$I2(u,v) = \text{sum}(I1(u-1:u+1, v-1:v+1).*H);$$

et $I2(u,v) = I1(u,v)$ pour les autres valeurs de (u,v) .

1. Expliquez pourquoi le filtre H ne modifie pas la luminosité globale de l'image.
2. Afin d'étudier le comportement de ce filtre sur un contour de l'image, on s'intéresse à une image élémentaire $I1$ qui présente une variation d'intensité sur quelques pixels dans la direction horizontale ; au sein d'une image plus complexe, cette image élémentaire pourrait être le bord d'un objet.

$I1 =$

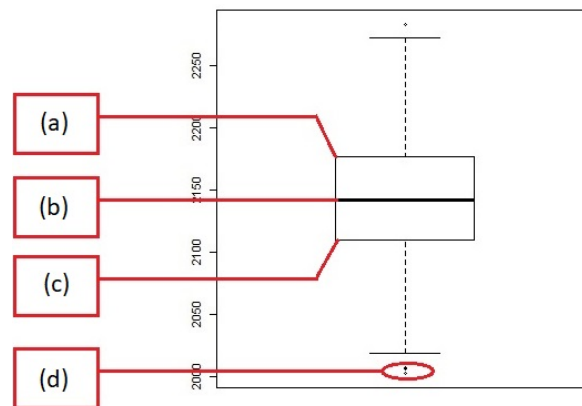
50.	50.	50.	50.	100.	150.	200.	200.	200.	200.
50.	50.	50.	50.	100.	150.	200.	200.	200.	200.
50.	50.	50.	50.	100.	150.	200.	200.	200.	200.
50.	50.	50.	50.	100.	150.	200.	200.	200.	200.
50.	50.	50.	50.	100.	150.	200.	200.	200.	200.

Calculez l'image $I2 = \text{conv}(I1, H)$.

3. En vous appuyant sur cet exemple, justifiez pourquoi on qualifie ce filtre de *réhausseur de contour*.
4. Quel serait le comportement de ce filtre vis-à-vis du bruit de l'image ?

Exercice 3 : Questions de cours (4 points)

1. Qu'est ce qu'une variable aléatoire ? (1 point)
2. Donner la différence entre un paramètre empirique et un paramètre théorique. (1 point)
3. Comment s'appelle de type de représentation graphique ? Donner les définitions des points (a), (b), (c) et (d). (1 point)



4. Que représente l'histogramme d'un échantillon d'un variable aléatoire ? Esquissez l'allure d'un histogramme. (1 point)

Exercice 4 : Estimation (9 points) Pour cet exercice, vous disposez d'une série de commandes R. Une partie du code est superflue et pour répondre aux questions suivantes, vous devez sélectionner uniquement les lignes de codes utiles. Les lignes sont numérotées, pour répondre à une question, citer les lignes de codes utilisées.

Attention : chaque ligne de code superflue donnée pour répondre à une question enlève 0.5 point.

On réalise 100 solutions d'eau chlorée dans les mêmes conditions expérimentales. On dispose donc de 100 réplicats. On mesure pour chaque réplicat la concentration en chlore. On suppose que la concentration en chlore suit une loi normale d'écart-type connu 0.07 mg/l. A partir de ces mesures, on calcule la moyenne empirique qui vaut 5.21 mg/l.

```
1  n <- 100
2
3  xbar <- 5.21
4  xbar <- 5.21/n
5
6  sig <- 0.07
7  sig <- 0.07*0.07
8  sig <- 0.07*n/(n-1)
9
10 c(xbar-qnorm(0.05)*sig/sqrt(n), xbar+qnorm(0.95)*sig/sqrt(n))
11 c(xbar-qnorm(0.025)*sig/sqrt(n), xbar+qnorm(0.975)*sig/sqrt(n))
12 c(xbar-qnorm(0.01)*sig/sqrt(n), xbar+qnorm(0.99)*sig/sqrt(n))
13
14 xbar+c(-1,1)*qnorm(0.95)*sig/sqrt(n)
15 xbar+c(-1,1)*qnorm(0.975)*sig/sqrt(n)
```

```

16 xbar+c(-1,1)*qnorm(0.99)*sig/sqrt(n)
17 xbar+c(-1,1)*qnorm(0.995)*sig/sqrt(n)
18
19 xbar+c(-1,1)*qt(n-1,0.995)*sig/sqrt(n)
20 xbar+c(-1,1)*qt(n-1,0.975)*sig/sqrt(n)
21
22 xbar+c(-1,1)*qchisq(n-1,0.95)*sig/sqrt(n)
23 xbar+c(-1,1)*qchisq(n-1,0.99)*sig/sqrt(n)
24
25 (qnorm(0.95)*sig^ 2/0.01^2)
26 (qnorm(0.95)*sig/0.01)^2
27 (qnorm(0.975)*sig^ 2/0.01^2)
28 (qnorm(0.975)*sig/0.01)^2
29
30 (qt(n-1,0.975)*sig/0.01)^2
31 (qt(n-1,0.95)*sig/0.01)^2
32
33 (qchisq(n-1,0.975)*sig/0.01^2)
34 (qchisq(n-1,0.95)*sig/0.01^2)
35
36 pnorm(0.01/sig^ 2*n)
37 pnorm(0.01/sig^ 2*sqrt(n))
38 pnorm(0.01/sig*sqrt(n))
39
40 alpha<-(1-pnorm(0.01/sig*sqrt(n)))/2
41 alpha<-2*(1-pnorm(0.01/sig*sqrt(n)))
42 alpha<-2*(1-pnorm(0.01/sig*sqrt(n)))
43 alpha<-1-pnorm(0.01/sig^ 2*sqrt(n))
44 alpha<-pnorm(0.01/sig*sqrt(n))
45
46 1-alpha
47 1-alpha/2
48 (1-alpha)/2

```

1. Rappeler l'objectif principal de l'estimation de paramètres. *(1 point)*
2. Donner les deux types d'estimations possibles. *(1 point)*
3. Parmi les lignes de code, choisir celles qui vous permettent de calculer l'intervalle de confiance à 95% de la vraie concentration de la solution. Donner le nom du quantile utilisé dans cet intervalle et justifier ce choix de quantile. *(3 points)*
4. Parmi les lignes de code, choisir celles qui permettent de calculer la taille de l'échantillon qui aboutit à un intervalle de niveau de confiance de 95% égal à 5.21 ± 0.01 . *(1 point)*
5. Parmi les lignes de code, choisir celles qui permettent de calculer la valeur de α telle que l'intervalle vaille 5.21 ± 0.01 . Quelle commande donne le niveau de confiance correspondant ? *(1 point)*
6. Donner la formule de l'intervalle de confiance pour la moyenne dans le cas gaussien, avec variance inconnue et nommer toutes les notations utilisées. *(2 points)*

Exercice 5 : Tests statistiques *(7 points)*

Pour cet exercice, vous disposez d'une série de commandes R. Une partie du code est superflue et pour répondre aux questions suivantes, vous devez donc sélectionner les lignes de codes utiles. Les lignes sont numérotées, pour répondre à une question, merci de citer les lignes de codes utilisées.

Attention : chaque ligne de code superflue donnée pour répondre à une question

enlève 0.25 point.

On souhaite étudier le nombre moyen d'enfants par femme en Isère. On étudie donc la variable nombre d'enfants sur une population de 300 personnes. Les données de cette variable ont été extraites et stockées dans la variable P. La variable P_1 contient les données concernant 100 femmes habitant des villes de plus de 100 000 habitant-es et P_2 contient celles concernant 200 femmes habitant des villes de moins de 100 000 habitant-es.

```
1 table(P_1)
2 mean(P_1)
3 var(P_1)
4 sd(P_1)
5
6 n<-length(P_1)
7
8 hist(P_1, prob=T)
9 hist(table(P_1), prob=T)
10
11 boxplot(P_1)
12 barplot(table(P_1))
13
14 t.test(P_1,mu=2, alternative="less")
15 t.test(P_1,mu=-2, alternative="less")
16 t.test(P_1,mu=2, alternative="greater")
17 t.test(P_1,mu=2)
18 prop.test(P_1,mu=2, alternative="less")
19 prop.test(P_1,mu=2, alternative="greater")
20 prop.test(P_1,mu=2)
21
22 t.test(P_1,P_2)
23 t.test(P_1, P_2, var.equal=FALSE)
24 t.test(P_1, P_2, var.equal=TRUE)
25 t.test(P_1, P_2, alternative = "less")
26 t.test(P_1, P_2, alternative="greater")
27 t.test(P_1, P_2, var.equal=TRUE, alternative = "less")
28 t.test(P_1, P_2, var.equal=FALSE, alternative="greater")
```

1. Indiquer les lignes de code qui permettent de calculer les statistiques descriptives de P_1 (les nommer) et de tracer sa distribution empirique. (1.5 point)
2. On souhaite tester le nombre moyen d'enfants par femme : est-il égal à 2 ou est-il plus petit ? Donner les hypothèses nulle et alternative \mathcal{H}_0 et \mathcal{H}_1 correspondant à ce test. (1 point)
3. Indiquer les lignes de code qui mettent en oeuvre le test de la question précédente. (1 point)
4. La p-valeur obtenue est de 0.40. Commenter. (1.5 point)
5. On veut savoir si le nombre moyen d'enfants par femme est différent entre les villes de plus de 100 000 habitant-es et celles de moins de 100 000 habitant-es. Définir les hypothèses nulle et alternative \mathcal{H}_0 et \mathcal{H}_1 pour réaliser ce test. (1 point)
6. Quel est le cadre de réalisation du test ? (0.5 point) Choisir les lignes de codes permettant de mettre en oeuvre le test. (0.5 point)