Exercices

Rappels, vecteurs gaussiens

Exercice 1. On donne les poids à la naissance (en kg) de 10 enfants :

$$3.2; 2.4; 3.3; 3.4; 3.9; 2.9; 3.3; 4.5; 1.9; 3.3.$$

1. Calculer le poids moyen, puis l'écart-type du poids des enfants.
2. On a demandé aux mères si elles avaient travaillé au cours de leur grossesse : cinq ont travaillé (n°1,3,4,7,10) et cinq n'avaient pas d'activité professionnelle (n°2,5,6,8,9). On calcule, pour chaque groupe, la moyenne et la variance du poids de l'enfant, et on obtient :

$$\begin{array}{rlll} \overline{y}_1 = 3.3 & \widehat{\sigma}_1(y) = 0.063 & min = 3.2 & max = 3.4 \\ \overline{y}_2 = 3.12 & \widehat{\sigma}_2(y) = 0.956 & min = 1.9 & max = 4.5 \end{array}$$

   Que constatez-vous ?
3. En supposant que l'on ne dispose que des effectifs (n_1 et n_2), moyennes ($\overline{y}_1$ et $\overline{y}_2$) et variances ($\widehat{\sigma}_1^2(y)$ et $\widehat{\sigma}_2^2(y)$) de chaque groupe, calculer la moyenne et l'écart-type des deux groupes réunis.

Exercice 2. On donne la matrice de variance-covariances empiriques Γ de trois variables y, x^1, x^2 :

$$\Gamma = \frac{1}{8}\begin{pmatrix} 1078 & 269 & 2 \\ 269 & 80 & 17 \\ 2 & 17 & 22 \end{pmatrix}$$

Calculer $var(x^1 + x^2)$, $r(x^1, x^2)$, $cov(3y, 2x^2 + 5\mathbb{1})$ et $r(3y, 2x^2 + 5\mathbb{1})$.

Exercice 3. Soit A une matrice réelle de taille $k \times d$ et soit X un vecteur aléatoire à valeurs dans $\mathbb{R}^d$, d'espérance nulle et de matrice de covariance Γ. Montrer que

$$\mathbb{E}\left(\left|AX\right|^2\right) = \mathrm{Tr}(A\Gamma A').$$

Spécifier la valeur obtenue dans le cas où $\Gamma = \mathrm{I}_d$, $k = d$ et A correspond à la projection orthogonale sur un sous-espace vectoriel $S \subset \mathbb{R}^d$

Exercice 4. Soit X un vecteur aléatoire gaussien à valeurs dans $\mathbb{R}^n$ de loi $\mathcal{N}_n(0, \Sigma)$.

1. Soit M une matrice orthogonale de taille n. Quelle est la loi du vecteur aléatoire MX ?
2. Soient u et v deux vecteurs de $\mathbb{R}^n$. Calculer la covariance $\mathrm{cov}\big(\langle X, u\rangle, \langle X, v\rangle\big)$.

Exercice 5. Pour $1 \le i \le N$, soit $v_i \in \mathbb{R}^d$ et g_i une variable réelle de loi $\mathcal{N}(0, \sigma_i^2)$. On suppose que les g_i sont indépendantes. Déterminer la loi du vecteur aléatoire

$$\sum_{i=1}^{N} g_i\, v_i.$$

Exercice 6. On considère pour $t \in \mathbb{R}$ la quantité

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{itx} e^{-x^2/2}\, dx.$$

Calculer $\varphi(0)$. Exprimer $\varphi'(t)$ en fonction de $\varphi(t)$ et en déduire la valeur de $\varphi(t)$.

Exercice 7. Soit X une variable aléatoire réelle de loi $\mathcal{N}(0,1)$ et ε une variable indépendante de X telle que $P(\varepsilon = 1) = P(\varepsilon = -1) = \frac{1}{2}$. Montrer que εX suit une loi $\mathcal{N}(0,1)$ mais que le vecteur $(X, \varepsilon X)$ n'est pas gaussien. Indication : calculer $P(X + \varepsilon X \neq 0)$.

Exercice 8. On considère une variable d'intérêt X de loi $\mathcal{N}(0,4)$. On note $\overline{X}_n$ la moyenne empirique dans le cadre du modèle d'échantillonnage. Calculer la probabilité que $\overline{X}_{10}$ soit supérieure à 1.04.

Exercice 9. On se situe dans le cadre où la variable d'intérêt X suit une loi Normale. On s'intéresse maintenant à la distribution de la variance empirique $\frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2$.
Les observations obtenues à l'issue d'une expérimentation sont les suivantes :

$$-2,\ 2,\ 4,\ 2,\ -1,\ 6,\ 3,\ 2,\ -1,\ 5$$

1. Calculer la moyenne $\overline{x}_n$ et la variance observée $\widehat{\sigma}^2(x) = \frac{1}{n}\sum_{i=1}^n (x_i - \overline{x}_n)^2$.
2. Quelle est la loi suivie par la v.a. $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2$?
3. $\widehat{\sigma}^2$ est-il un estimateur sans biais de σ^2 ?
4. On pense que la variance supposée de X dans la population vaut 4, mais on se doute qu'elle est peut-être sous-estimée. Pour répondre à cette interrogation, on calcule la probabilité que $\widehat{\sigma}^2$ dépasse la valeur estimée sur l'échantillon $\widehat{\sigma}^2(x)$, sachant que la valeur théorique est $\sigma^2 = 4$. Commentez.

Exercice 10. On souhaite tester l'efficacité d'un régime alimentaire censé faire perdre du poids. On dispose des observations suivantes :

i	1	2	3	4	5	6	7	8	9	10	11	12
y_i	72	84	85	97	100	76	80	80	95	100	70	80
z_i	70	74	80	100	80	85	65	88	91	89	77	86

où y_i désigne le poids de l'individu i avant d'avoir commencé le régime, et z_i le poids de ce même individu après avoir suivi le régime. On suppose que les $X_i = Z_i - Y_i$ sont *i.i.d* et suivent une loi normale.

1. Calculer la perte de poids moyenne, et l'intervalle de confiance de niveau de confiance $1 - \alpha$.
2. Formuler les hypothèses H_0 et H_1 du test permettant de juger de l'efficacité du régime. Indiquer quelle est votre statistique de test ? Donner la région de rejet de H_0 (on prendra $\alpha = 5\%$). Conclure.

Modèle linéaire, généralités

Exercice 11. On considère les modèles suivants :

1. $Y_i = \theta_1 + \theta_2 x_i + E_i\,, \quad i = 1, \ldots, n,$
2. $Y_i = \theta + E_i\,, \quad i = 1, \ldots, n,$
3. $Y_i = \theta_0 + \theta_1(x_i - \overline{x}) + E_i\,, \quad i = 1, \ldots, n,$
4. $Y_i = \theta_0 + \theta_1 x_i^1 + \theta_2 x_i^2 + E_i\,, \quad i = 1, \ldots, n.$

Dans chaque cas, donner le vecteur θ et la matrice X pour écrire le modèle proposé sous la forme $Y = X\theta + E$, sachant que Y et E sont des vecteurs de $\mathbb{R}^n$, constitués respectivement des éléments $(Y_i)_{i=1,\ldots,n}$ et $(E_i)_{i=1,\ldots,n}$.

Exercice 12. On a observé deux variables $\mathbf{x}$ et $\mathbf{y}$ sur un échantillon de 9 individus. Les résultats sont les suivants :

x_i	10	13	14	16	19	20	21	24	25
y_i	32	37	35	40	46	45	42	48	53

On souhaite réaliser la régression linéaire simple de $\mathbf{y}$ sur $\mathbf{x}$:

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

où e_i est la réalisation d'une v.a. $E_i \sim \mathcal{N}(0, \sigma^2)$, avec E_i indépendantes et σ^2 inconnue.

1. Ecrire le modèle de régression linéaire de Y sur X sous forme matricielle, en spécifiant les matrices et vecteurs.
2. Calculer $X'X$, $X'y$ et $(X'X)^{-1}X'y$. Donner les cœfficients de la régression.
3. Calculer le vecteur des valeurs ajustées $\widehat{y}$ et le vecteur des résidus estimés $\widehat{e}$. Vérifier que la somme des résidus estimés vaut 0. En déduire l'estimation de σ^2.

Note : on vous donne ci-dessous quelques résultats intermédiaires utiles

$$\sum_{i=1}^{n} x_i = 162 \quad \sum_{i=1}^{n} y_i = 378 \quad \sum_{i=1}^{n} x_i^2 = 3124 \quad \sum_{i=1}^{n} x_i y_i = 7064$$

$$\begin{pmatrix} 9 & 162 \\ 162 & 3124 \end{pmatrix}^{-1} = \begin{pmatrix} 3124/1872 & -18/208 \\ -18/208 & 1/208 \end{pmatrix}$$

Exercice 13. Soient $Y \in \mathbb{R}^n$ et $X \in M_{n,p}(\mathbb{R})$ une matrice de rang p. On considère la fonction $\varphi : \mathbb{R}^p \to \mathbb{R}$ définie par $\varphi(\theta) = |Y - X\theta|^2$. Différentier deux fois la fonction φ. En déduire une preuve analytique de l'expression de l'estimateur des moindres carrés.

Exercice 14. Soient $Y \in \mathbb{R}^n$ et $X \in M_{n,p}(\mathbb{R})$. Soit $\lambda > 0$. On pose

$$\psi(\theta) = |Y - X\theta|^2 + \lambda|\theta|^2.$$

1. Montrer que ψ est strictement convexe et admet un unique minimum (même si X n'est pas injective).
2. Pour quelle valeur de θ le minimum de ψ est-il atteint ?

Exercice 15. On considère le modèle $y = X\theta + e$ où :

- y et e sont des vecteurs de $\mathbb{R}^n$ (avec $n = 50$),
- e est une réalisation de $E \sim \mathcal{N}_n(0, \sigma^2 I_n)$,
- θ est le vecteur de $\mathbb{R}^3$ des paramètres à estimer,

– X est une matrice $M(n,3)$.

On lui associe les résultats matriciels suivants :

$$|y| = 38 \qquad X'X = \begin{pmatrix} 50 & 0 & 2 \\ 0 & 20 & -5 \\ 2 & -5 & 4 \end{pmatrix} \qquad X'y = \begin{pmatrix} 250 \\ 40 \\ 10 \end{pmatrix}$$

$$(X'X)^{-1} = \begin{pmatrix} 0.021 & -0.004 & -0.015 \\ -0.004 & 0.073 & 0.094 \\ -0.015 & 0.094 & 0.375 \end{pmatrix}$$

1. Préciser les lois des estimateurs $\widehat{\theta}$ de θ, et $\widehat{\sigma}^2$ de σ^2.
2. Calculer les résultats numériques associés à ces estimateurs.
3. Donner la distribution de $\widehat{\theta}_2$.
4. Calculer un intervalle de confiance de niveau 95% pour θ_2. Quelle conclusion pouvez-vous en tirer ?

Exercice 16. On note λ la mesure de Lebesgue sur $\mathbb{R}$.

1. Soit $\mu(dx) = f(x)\,dx$ une loi de probabilité sur $\mathbb{R}$. Soit $a \in \mathbb{R}^+$ et $A = \{x \in \mathbb{R};\ f(x) \geq a\}$. Soit B un ensemble mesurable tel que $\mu(A) = \mu(B)$. Démontrer que
$$\lambda(A) \leq \lambda(B).$$
Indication : on pourra commencer par traiter le cas particulier où $A \cap B = \emptyset$.
2. Trouver un ensemble de probabilité $1-\alpha$ pour la loi de student $s(\nu)$ et de mesure de Lebesgue la plus petite possible. Quel est le rapport avec les intervalles de confiance ?

Exercice 17. On considère le modèle linéaire gaussien $Y = X\theta + \sigma G$.

1. Déterminer la loi de $\widehat{Y}$.
2. En déduire que $\mathbb{E}\Big(\big|\widehat{Y} - X\theta\big|^2\Big) = p\sigma^2$

Exercice 18. On considère le modèle linéaire gaussien $Y = X\theta + \sigma G$, avec X de taille $n \times p$ et de rang p.

1. Quelle est la loi de Y ? Exprimer la densité de cette loi par rapport à la mesure de Lebesgue.
2. On veut estimer $(\theta, \sigma) \in \mathbb{R}^p \times \mathbb{R}_+$ à partir de l'observation de Y. Montrer que le modèle est dominé. Exprimer la log-vraisemblance.
3. Calculer l'estimateur du maximum de vraisemblance de (θ, σ). Comparer avec l'estimateur obtenu par la méthode des moindres carrés.

Exercice 19. On considère une régression linéaire simple dans sa forme générale $Y_i = \theta_0 + \theta_1 x_i + \epsilon_i$, $i = 1, \ldots, n$. Sa formulation matricielle est $Y = X\theta + E$. Donner une base orthonormée de l'image de X. En déduire l'expression de $\widehat{Y} = P_{\mathrm{Im}(X)} Y$. Retrouver ainsi les formules pour $\widehat{\theta}$.

Exercice 20. On considère un modèle M de régression linéaire simple :

$$y_i = \theta_0 + \theta_1 x_i + e_i, \quad i = 1, \ldots, n,$$

ainsi que le sous modèle M_0 : $\qquad y_i = \beta_0 + \varepsilon_i, \quad i = 1, \ldots, n.$

1. Exprimer $\widehat{\beta}_0$ et $\widehat{\theta}_0$ en fonction des observations x et y.
2. Donner une condition nécessaire et suffisante pour que $\widehat{\beta}_0 = \widehat{\theta}_0$.

3. Proposer une condition, portant seulement sur le prédicteur x, qui assure l'égalité $\widehat{\beta}_0 = \widehat{\theta}_0$, quelle que soit la valeur de y.
4. Généraliser le résultat de la question précédente au cas où le modèle M est une régression linéaire multiple faisant intervenir des prédicteurs $x^1, x^2, \ldots, x^{p-1}$.

Exercice 21. Soient $n \geq p \geq p_0$. On considère une modèle linéaire $Y = X\theta + E$ où X est de taille $n \times p$ et s'écrit par blocs $X = (X_0 \ X_1)$ où X_0 est de taille $n \times p_0$. On considère le sous-modèle $Y = X_0\theta_0 + E$ qui cherche à expliquer Y en fonction des p_0 premiers prédicteurs seulement. On suppose que X et X_0 sont de rang plein.

1. Montrer que $\widehat{Y_0} = P_{\mathrm{Im}(X_0)}\widehat{Y}$ et en déduire que $\widehat{\theta_0} = (X_0'X_0)^{-1}X_0'X\widehat{\theta}$.
2. En exprimant X par blocs, en déduire que
$$\widehat{\theta_0} = P_0\widehat{\theta} + (X_0'X_0)^{-1}X_0'X_1P_1\widehat{\theta},$$
où les projections P_0, P_1 sont définies pour tout $z \in \mathbb{R}^p$ par $P_0 z = (z_1, \ldots, z_{p_0})$ et $P_1 z = (z_{p_0+1}, \ldots, z_p)$.
3. Montrer que si $\mathrm{Im}(X_0) \perp \mathrm{Im}(X_1)$ alors $\widehat{\theta_0} = P_0\widehat{\theta}$. Interprétez ce résultat.
4. On suppose maintenant que $(X_0'X_0)^{-1}X_0'X_1 = 0$. En déduire que $P_{\mathrm{Im}(X_0)}X_1 = 0$ puis que $\mathrm{Im}(X_0) \perp \mathrm{Im}(X_1)$.

Exercice 22. Lorque l'on observe deux variables $\mathbf{x}$ et $\mathbf{y}$ sur un échantillon de taille n, on peut réaliser deux régressions linéaires :

$$y_i = \theta_0 + \theta_1 x_i + e_i, \quad i = 1, \ldots, n \qquad (R1)$$
$$x_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \ldots, n \qquad (R2)$$

On notera D_1 et D_2 les droites de régressions associées.

1. Montrer que $(\overline{x}, \overline{y}) \in D_1 \cap D_2$.
2. Montrer que $\widehat{\theta}_1\widehat{\beta}_1 = \big(r(x,y)\big)^2$. Les deux droites de régression coincident-elles ?
3. Dans ce qui suit, on construit une droite qui approche les données de manière plus symétrique en les variables $\mathbf{x}, \mathbf{y}$. Pour toute droite $\Delta \subset \mathbb{R}^2$, on définit
$$F(\Delta) := \frac{1}{n}\sum_{i=1}^{n} \mathrm{dist}\big((x_i, y_i), \Delta\big)^2.$$
Pour simplifier on suppose que $\overline{x} = \overline{y} = 0$ et on ne considère que les droites passant par l'orgine.
   (a) Montrer que si Δ a pour équation $ax + by = 0$ avec $a^2 + b^2 = 1$ alors
   $$F(\Delta) = a^2 var(\mathbf{x}) + b^2 var(\mathbf{y}) + 2ab\, cov(\mathbf{x}, \mathbf{y}) = \left\langle \begin{pmatrix} a \\ b \end{pmatrix}, M \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle,$$
   où $M = \begin{pmatrix} var(\mathbf{x}) & cov(\mathbf{x}, \mathbf{y}) \\ cov(\mathbf{x}, \mathbf{y}) & var(\mathbf{y}) \end{pmatrix}$ est la matrice de covariance empirique de $(\mathbf{x}, \mathbf{y})$.
   (b) En vous aidant de la diagonalisation de M en base orthonormée, déterminer la ou les droites Δ passant par l'origine qui minimisent la fonctionnelle F.
   (c) Interpréter le résultat obtenu.

Exercice 23. On considère le modèle linéaire $Y = X\theta + E$ (avec p paramètres et n individus) et $\widehat{\theta} = (X'X)^{-1}X'Y$ l'estimateur des moindres carrés. Soit $B \in M_{p,n}(\mathbb{R})$. On suppose ici que $\widetilde{\theta} := BY$ est un estimateur sans biais de θ.

1. Montrer que l'absence de biais de $\widetilde{\theta}$ se traduit par $BX = I_p$.
2. On pose $A = (X'X)^{-1}X'$ et $C = B - A$. Calculer CX. Exprimer BB' en fonction de A et C. En déduire que $BB' \geq AA'$.
3. Conclure que $\mathrm{Cov}\big(\widetilde{\theta}\big) \geq \mathrm{Cov}\big(\widehat{\theta}\big)$.

Régression linéaire, tests associés

Exercice 24.

1. Soit X une variable aléatoire réelle de loi de student $s(\nu)$. Quelle est la loi de X^2 ?
2. Si Z suit une loi de Fisher $F(n_1, n_2)$, quelle est la loi de $1/Z$?
3. Déterminer explicitement la loi de Fisher $F(n_1, n_2)$. Etudier le sens de variation de sa densité. On rappelle que la loi du $\chi^2(n)$ admet pour densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^+$ la fonction $t \mapsto e^{-\frac{t}{2}} t^{\frac{n}{2}-1} 2^{-\frac{n}{2}} / \Gamma(\frac{n}{2})$.

Exercice 25.

1. Montrer que lorsque n tend vers l'infini, la suite des lois de Student $\big(S(n)\big)_{n\geq 1}$ converge étroitement vers la loi $\mathcal{N}(0,1)$. Indication : on pourra utiliser la loi des grands nombres.
2. Etudier la limite lorsque n tend vers l'infini des suites de lois $\big(F(n_1, n)\big)_{n\geq 1}$ et $\big(F(n, n)\big)_{n\geq 1}$

Exercice 26. On considère le modèle linéaire gaussien $Y = X\theta + \sigma G$ avec $\theta \in \mathbb{R}^p$, X de taille $n \times p$ et de rang p et G de loi $\mathcal{N}_n(0, I_n)$. Soit $v \in \mathbb{R}^p$ un vecteur non nul. On note $\widehat{\theta}$ pour l'estimateur des moindres carrés.

1. Déterminer la loi de $\dfrac{\langle \widehat{\theta} - \theta, v \rangle}{\sigma}$, puis celle de $\dfrac{\langle \widehat{\theta} - \theta, v \rangle}{\widehat{\sigma}}$.
2. En déduire un intervalle de confiance pour $\langle \theta, v \rangle$.
3. Construire un test (de type Student) pour H_0 : "$\langle \theta, v \rangle = c$" contre H_1 : "$\langle \theta, v \rangle \neq c$".
4. Expliciter le test obtenu pour l'hypothèse nulle "$\theta_1 = \theta_2$".

Exercice 27. On appelle fréquence seuil d'un sportif amateur, sa fréquence cardiaque obtenue après trois quarts d'heure d'un effort soutenu de course à pied. Elle est mesurée à l'aide d'un cardio-fréquence-mètre. On cherche à savoir si l'âge d'un sportif a une influence sur sa fréquence seuil. On dispose des résultats suivants où x_i représente l'âge du sportif et y_i sa fréquence seuil.

x_i	30	54	29	51	36	41	40	23	49	30
y_i	175	165	169	172	170	170	167	170	166	167

x_i	32	22	22	32	44	34	32	20	46	45
y_i	177	169	172	173	168	169	170	172	175	168

On modélise y_i ($i = 1, ...n$ avec $n = 20$) en fonction de x_i en supposant un modèle de régression linéaire : $y_i = \beta_0 + \beta_1 x_i + e_i$ où e_i est la réalisation d'une v.a. E_i distribuée selon une loi $N(0, \sigma^2)$ avec σ^2 inconnue et E_i ($i = 1, ..., n$) indépendantes.

1. Estimer les paramètres inconnus β_0, β_1 et σ^2.

2. Proposer un intervalle de confiance à 95% pour β_1. Quelle première conclusion peut-on en tirer ?
3. Tester l'absence d'effet de l'âge du sportif sur sa fréquence seuil. Conclure.
4. Pour une nouvelle valeur $x_0 = 65$, calculer la prédiction naturelle $\widehat{y}_0$ de y_0 et trouver un intervalle de prédiction à 95% pour y_0.

On vous donne les résultats intermédiaires utiles pour les calculs :

$$\sum_{i=1}^{20} x_i = 712 \qquad \sum_{i=1}^{20} y_i = 3404 \qquad \sum_{i=1}^{20} x_i y_i = 120987$$

$$\sum_{i=1}^{20} (x_i)^2 = 27338 \qquad \sum_{i=1}^{20} (y_i)^2 = 579550 \qquad (X'X)^{-1} = \frac{1}{39816}\begin{pmatrix} 27338 & -712 \\ -712 & 20 \end{pmatrix}$$

Exercice 28. Un physicien est confronté au problème suivant : étudier l'influence de deux variables physiques quantitatives x^1 et x^2 sur une variable réponse y, caractérisant une matière première. Pour cela, il dispose de 10 échantillons de cette matière première pour lesquels les valeurs de x^1 et de x^2 sont connues et sur lesquels il mesure la variable y. Il a également calculé, pour chaque échantillon, la moyenne de x^1 et x^2, notée x^m. Les données sont les suivantes :

Numéro de l'échantillon	1	2	3	4	5	6	7	8	9	10
Valeur de x^1	2	1	3	4	5	1	2	3	5	4
Valeur de x^2	1	2	3	1	4	4	5	5	4	1
Valeur de y	6.2	2.9	5.7	10.3	9.9	0.8	2.4	5.2	8.1	6.6
Valeur de x^m	1.5	1.5	3.0	2.5	4.5	2.5	3.5	4.0	4.5	2.5

A partir de ces données, on a calculé les quantités suivantes :

$$\sum_{i=1}^{10} x_i^1 = 30 \qquad \sum_{i=1}^{10} x_i^2 = 30 \qquad \sum_{i=1}^{10} y_i = 58.1 \qquad \sum_{i=1}^{10} x_i^m = 30$$

$$\sum_{i=1}^{10} (x_i^1)^2 = 110 \qquad \sum_{i=1}^{10} (x_i^2)^2 = 114 \qquad \sum_{i=1}^{10} (y_i)^2 = 426.05 \qquad \sum_{i=1}^{10} (x_i^m)^2 = 101$$

$$\sum_{i=1}^{10} x_i^1 x_i^2 = 90 \qquad \sum_{i=1}^{10} x_i^1 y_i = 211.2 \qquad \sum_{i=1}^{10} x_i^2 y_i = 159.2$$

$$\sum_{i=1}^{10} x_i^1 x_i^m = 100 \qquad \sum_{i=1}^{10} x_i^2 x_i^m = 102 \qquad \sum_{i=1}^{10} y_i x_i^m = 185.2$$

1. Le physicien cherche donc à expliquer la variable réponse y en fonction de x^1 et x^2.
   (a) Ecrire le modèle de régression linéaire (noté $M1$) de y_i sur x_i^1 et x_i^2.
   (b) Construire la matrice d'incidence X et le vecteur des paramètres à estimer β tels que le modèle $(M1)$ puisse s'écrire sous la forme :
   $$y = X\beta + e$$
   (c) Calculer $\widehat{\beta}(y)$.
   On vous donne :
   $$\begin{pmatrix} 10 & 30 & 30 \\ 30 & 110 & 90 \\ 30 & 90 & 114 \end{pmatrix}^{-1} = \begin{pmatrix} 0.925 & -0.150 & -0.125 \\ -0.150 & 0.050 & 0.000 \\ -0.125 & 0.000 & 0.04167 \end{pmatrix}$$
   (d) Montrer que $SST = 88.489$ et $SSR = 10.83$.

(e) Tester l'hypothèse selon laquelle la variable y ne dépend ni de x^1 ni x^2.

(f) Tester l'hypothèse d'absence d'effet de x^1 sur y.

2. Dans un deuxième temps, le physicien propose d'utiliser la moyenne des deux variables x^1 et x^2, notée x^m, comme seule variable explicative. On considère donc le modèle de régression linéaire simple de y sur x^m :

$$(M2) : y_i = \alpha_0 + \alpha_1 x_i^m + e_i$$

(a) Estimer les paramètres de ce modèle.

(b) Tester l'absence d'effet de x^m sur y.

(c) Pour le modèle $(M1)$, on a obtenu $R_1^2 = 0.8767$ et pour le modèle $(M2)$, $R_2^2 = 0.1221$. Comparer les deux modèles à l'aide d'un critère approprié.

(d) Expliquer pourquoi l'idée d'utiliser x^m comme variable explicative au lieu de x^1 et x^2 est particulièrement inadaptée à ces données.

3. Le physicien pourrait-il estimer un modèle de régression linéaire multiple de y en fonction de x^1, x^2 et x^m ? Justifier votre réponse.

Exercice 29. L'association *Air Breizh* qui surveille la qualité de l'air en Bretagne mesure, depuis 1994, la concentration en ozone (en μg/ml) toutes les 10 minutes et obtient donc le maximum journalier de la concentration en ozone. Elle collecte également à certaines heures de la journée, des données météorologiques (comme la température, le vent, la nébulosité, ...) susceptibles d'avoir une influence sur la concentration en ozone. On cherche donc à modéliser la concentration en ozone en fonction de variables météorologiques.

Pour cela, nous traitons un jeu de 50 données journalières pour lesquels on dispose des variables suivantes :
- la concentration en ozone maximale (en μg/ml), notée y ;
- la nébulosité mesurée à midi, notée x^1 ;
- la température mesurée à midi, notée x^2 ;
- une variable synthétique représentant le vent et tenant compte de sa direction et de sa vitesse, notée x^3.

Sur les 50 observations disponibles, on a calculé les quantités suivantes :

$$\sum_{i=1}^{50} x_i^1 = 251 \qquad \sum_{i=1}^{50} x_i^2 = 1018 \qquad \sum_{i=1}^{50} x_i^3 = -42 \qquad \sum_{i=1}^{50} y_i = 4315$$

$$\sum_{i=1}^{50} (x_i^1)^2 = 1575 \qquad \sum_{i=1}^{50} x_i^1 x_i^2 = 4823 \qquad \sum_{i=1}^{50} x_i^1 x_i^3 = -978 \qquad \sum_{i=1}^{50} x_i^1 y_i = 19377$$

$$\sum_{i=1}^{50} (x_i^2)^2 = 21824 \qquad \sum_{i=1}^{50} x_i^2 x_i^3 = -657 \qquad \sum_{i=1}^{50} x_i^2 y_i = 90775$$

$$\sum_{i=1}^{50} (x_i^3)^2 = 9096 \qquad \sum_{i=1}^{50} x_i^3 y_i = 4833 \qquad \sum_{i=1}^{50} (y_i)^2 = 400337$$

1. Dans un premier temps, on souhaite modéliser la concentration en fonction d'une seule variable par une régression linéaire simple. On vous donne dans le tableau ci-dessous les

cœfficients de corrélation linéaire entre la concentration en ozone et chacune des trois variables météorologiques :

	x^1	x^2	x^3
y	-0.7698	0.5275	0.5314

(a) Au vu de ces résultats, laquelle des trois variables météorologiques permettrait d'obtenir la meilleure régression linéaire simple pour expliquer la concentration en ozone au sens du R^2 maximum ? Justifier ce choix.

(b) Donner l'indice R^2 de ce modèle.

(c) Calculer SST. En déduire SSL et SSR.

(d) Ecrire ce modèle de régression linéaire simple et estimer les paramètres correspondants.

2. Dans un deuxième temps, on considère la régression linéaire multiple contenant les trois variables météorologiques sous la forme :

$$y_i = \beta_0 + \beta_1 x_i^1 + \beta_2 x_i^2 + \beta_3 x_i^3 + e_i$$

où e_i est la réalisation de v.a. indépendantes E_i normales centrées, de variance σ^2 inconnue.

(a) Estimer les paramètres de ce modèle.
On vous donne la matrice $(X'X)^{-1}$ correspondant au modèle :

$$\begin{pmatrix} 50 & 251 & 1018 & -42 \\ 251 & 1575 & 4823 & -978 \\ 1018 & 4823 & 21824 & -657 \\ -42 & -978 & -657 & 9096 \end{pmatrix}^{-1} \approx \begin{pmatrix} 0.94548 & -0.05459 & -0.03215 & -0.00382 \\ -0.05459 & 0.00546 & 0.00135 & 0.00043 \\ -0.03215 & 0.00135 & 0.00125 & 0.00009 \\ -0.00382 & 0.00043 & 0.00009 & 0.00014 \end{pmatrix}$$

(b) Calculer la SSR (on trouvera $SSR \approx 8000$ et l'on réfléchira aux chiffres significatifs suivant les arrondis effectués). En déduire l'estimation de σ^2 et l'indice de qualité R^2.

(c) Tester l'absence d'effet de chaque variable explicative.

3. Comparer, par un test approprié, les deux modèles estimés et choisir le meilleur modèle. On vous demande d'expliquer précisément le test que vous utilisez (hypothèses, statistique utilisée et loi de cette statistique, degrés de liberté et valeurs lues dans la table).

4. Conclure et commenter l'ensemble des résultats du modèle sélectionné (ajustement du modèle, effet des variables, interprétation des cœfficients de régression, ...) en insistant sur les conditions météorologiques favorisant une pollution à l'ozone.

Analyse de variance

Exercice 30. Un sportif désire étudier, parmi trois régimes (codés de 1 à 3), celui qui est le mieux adapté pour la course. Il s'est chronométré 9 jours de suite sur un parcours donné et a noté à chaque fois le régime suivi. Il a obtenu les résultats suivants :

Régime	Temps		
1	44.5	43.0	41.5
2	40.0	42.0	41.0
3	40.0	44.0	42.0

On note $\mathbf{y}$ la variable Temps, et y_{ij} est la valeur observée de la variable $\mathbf{y}$ au jour j ($j = 1, ..., n_i$) du régime i ($i = 1, 2, 3$). On considère un modèle à un facteur (à 3 niveaux) :

$$y_{ij} = \mu_i + e_{ij} \text{ avec } i = 1, 2, 3 \text{ et } j = 1, ..., n_i$$

où les e_{ij} sont des réalisations indépendantes de E_{ij} distribuées selon une loi $\mathcal{N}(0, \sigma^2)$, avec σ^2 inconnue.

1. Donner les caractéristiques du plan d'expérience : variable réponse, facteur, nombre de niveaux, nombres d'observations par niveau, propriétés du plan (complet, équilibré ou répété ?).
2. Estimer les paramètres μ_i du modèle écrit ci-dessus. Calculer les variances du temps de course par régime. Commentez ces premières résultats.
3. Ecrire le modèle sous la forme de la paramétrisation centrée. Estimer les paramètres intervenant dans cette paramétrisation.
4. Déterminer la somme des carrés des résidus (SSR). En déduire l'estimation $\widehat{\sigma}^2(y)$ de σ^2.
5. Calculer la somme des carrés expliquée par le modèle (SSL). En déduire le R^2.
6. Tester l'absence d'effet du régime sur le temps de course, en formulant les hypothèses nulle et alternative, en calculant la statistique de test de Fisher et en concluant au risque 5%.

Exercice 31. On a relevé la production de lait (en hectolitres) pendant une année pour des vaches de 7 exploitations différentes des monts d'Auvergne. Les résultats sont résumés dans le tableau suivant :

Exploitation	1	2	3	4	5	6	7
Nombre de vaches	9	7	7	6	7	6	6
Production moyenne	26.88	22.34	19.54	18.95	27.17	25.87	25.72
Variance de la production	13.54	18.59	19.87	28.42	17.17	10.56	2.64

Soit y_{ij} la production de lait la j^e vache dans l'exploitation i, avec $i = 1, ..., 7$ et $j = 1, ..., n_i$. On pose le modèle :

$$y_{ij} = \mu_i + e_{ij}$$

où les e_{ij} sont des réalisations indépendantes d'une v.a. de loi $\mathcal{N}(0, \sigma^2)$.

1. Donner, sans calcul, les paramètres estimés du modèle.
2. Estimer le paramètre σ^2.
3. Tester l'hypothèse d'égalité des productions moyennes des 7 exploitations.
4. Donner un intervalle de confiance de $\mu_4 - \mu_5$ de niveau de sécurité 95%.

Exercice 32. Des patients atteints d'hypercholestérolémie sont suivis dans un hôpital. Pour réduire leur taux de cholestérol, on leur a proposé trois médicaments différents et/ou la mise en place d'un régime alimentaire hypercholestérolémiant. Les patients ont donc le choix entre six traitements. Après quelques mois de traitement, on a mesuré la baisse du taux de cholestérol et on a obtenu les résultats suivants :

	Médicaments		
Régime alimentaire	Aucun	A	B
Inchangé	0.2 0.4	1.0 1.2 0.9	1.0 0.9
Hypercholestérolémiant	0.8 0.6 0.7	1.0 0.9 1.1 1.2	1.5 1.6

On considère le modèle à deux facteurs avec interactions :

$$y_{ijl} = \mu_{ij} + e_{ijl}$$

où $i = \{1, 2\}$ est l'indice du facteur *"Régime"*,
$j = \{1, 2, 3\}$ correspond au niveau du facteur *"Médicament"*,
$l = \{1, ..., n_{ij}\}$ correspond à l'indice de l'observation dans la cellule (i, j),
et e_{ijl} est la réalisation d'une E_{ijl} distribuée selon une loi $N(0, \sigma^2)$ avec E_{ijl} indépendantes.

1. Quelles sont les caractérisques du plan d'expérience ?
2. Estimer les paramètres μ_{ij} du modèle.
3. Tracer le graphique des interactions.
4. Calculer SST, SSL et SSR. En déduire le coefficient de détermination R^2.
5. Tester l'absence d'interactions entre les deux facteurs. Pour cela, on vous donne ci-après les résultats concernant le modèle sans interactions.

```
                                     Sum of
Source                   DF         Squares
Model                     3      1.57231752
Error                    12      0.38518248
Corrected Total          15      1.95750000
```

6. Commenter les résultats du modèle de façon complète (en faisant intervenir tous les éléments que vous avez à votre disposition).

Exercice 33. Dans cet exercice, on cherche à étudier si certaines caractéristiques pourraient avoir un effet éventuel sur le temps de trajet domicile-travail le matin. Pour cela, nous modélisons le temps de trajet (en minutes) en fonction de deux variables qualitatives :
- le mode de transport à 2 modalités : voiture ($i = 1$) ou transport en commun ($i = 2$) ;
- le sens du trajet à 2 modalités : du centre ville à la banlieue (j=1) ou de la banlieue au centre ville (j=2).

Les résultats de l'expérience portant sur 40 observations au total, sont présentés dans le tableau suivant en terme de nombre d'observations, moyenne et variance du temps de trajet (en minutes),
- pour chaque combinaison possible entre le mode de transport et le sens du trajet,
- pour chaque niveau du mode de transport,
- pour chaque niveau du sens du trajet,
- et pour toutes les observations.

Nombre d'observations Moyenne Variance	Sens du trajet		Tous Sens confondus
Mode de transport	CV vers B ($j = 1$)	B vers CV ($j = 2$)	
Voiture ($i = 1$)	10 19.70 14.61	10 38.30 30.01	20 29.00 108.8
Transport en commun ($i = 2$)	10 27.50 19.05	10 26.90 14.89	20 27.20 17.06
Tous modes confondus	20 23.60 32.04	20 32.60 54.94	40 28.10 63.74

On note y_{ijl}, le l^e ($l = 1, ...10$) temps de trajet observé avec un mode de transport i ($i = 1, 2$) et un sens de trajet j ($j = 1, 2$). On considère un modèle à deux facteurs croisés avec interactions sous la forme :

$$y_{ijl} = \mu + \alpha_i^L + \alpha_j^C + \alpha_{ij} + e_{ijl} \text{ pour } i = 1, 2,\ j = 1, 2 \text{ et } l = 1, ..., 10.$$

1. Spécifier les caractéristiques du plan d'expérience.
2. Après avoir posé les contraintes associées à la paramétrisation proposée, estimer les paramètres inconnus μ, α_i^L, α_j^C et α_{ij} (pour $i = 1, 2$ et $j = 1, 2$).
3. Tracer le diagramme d'interactions.
4. Calculer la somme des carrés des résidus (SSR) et la somme des carrés totale (SST). En déduire SSL et R^2.
5. Chercher un sous-modèle pertinent en testant l'absence d'interaction ou d'effet des facteurs.
6. Commenter l'ensemble des résultats (diagramme des interactions, estimations, résultats des tests, ...) en insistant sur les caractéristiques favorisant une réduction ou au contraire une augmentation du temps de trajet.

Exercice 34. On injecte à 24 lapins de l'insuline en leur donnant des doses notées D_1, D_2 et D_3, préparées suivant deux protocoles différents notés P_1 et P_2. La réduction de sucre dans leur sang a été mesurée et elle a donné les résultats suivants :

Réduction	P_1	P_2
D_1	17 21 49 54	33 37 40 16
D_2	64 48 34 63	41 64 34 64
D_3	62 72 61 91	56 62 57 72

On note $y_{i,j,k}$ la réduction de sucre dans le sang du k^{eme} lapin ayant reçu une dose D_i d'insuline préparée selon le protocole P_j, avec $i = 1, 2, 3$ et $j = 1, 2$. On considère le modèle à deux facteurs

$$y_{i,j,k} = m_{i,j} + e_{i,j,k} \text{ pour } i = 1, 2, 3,\ j = 1, 2 \text{ et } k = 1, 2, 3, 4$$

où les $e_{i,j,k}$ sont les réalisations indépendantes d'une v.a. distribuée selon une loi $\mathcal{N}(0, \sigma^2)$. On s'intéresse aussi à la paramétrisation centrée :

$$y_{i,j,k} = \mu + a_i + b_j + c_{i,j} + e_{i,j,k} \text{ pour } i = 1, 2, 3,\ j = 1, 2 \text{ et } k = 1, 2, 3, 4.$$

1. A l'aide d'un tableur, estimer tous les paramètres inconnus $(m_{i,j}, \mu, a_i, b_j, c_{i,j})$ et calculer toutes les sommes de carrés associées (SST,SSR,SSL,SS1,SS2,SSI). On pourra construire un tableau sur le modèle de l'exercice précédent et le compléter.

   Indication : exemples de syntaxe sous OpenOffice : pour calculer la moyenne des contenus des cases $B7, C7, D7, E, 7, G1, H1, I1$, puis la variance empirique des contenus des cases $B12, C12, D12$, on utilise les commandes

   `=MOYENNE(B7:E7;G1:I1)`   `=VAR.P(B12:D12)`

   On tirera aussi partie de la fonction copier-coller et de la possibilité de décrire les cases de manière relative ou absolue (dans une formule de la case $A11$, l'expression $B12$ signifie "la case située une colonne plus à droite et une ligne plus bas", alors que $\$B12$ signifie "la case de la colonne B située une ligne plus bas").
2. Tracer le diagramme d'interactions. Commenter.
3. Tester l'hypothèse d'absence d'interactions entre les deux facteurs.

Table de quantiles des lois de Student S(d)

Donne le quantile t tel que S(d)(]-∞,t])=p

Nombre de degrés de liberté: d	Probabilité: p										
	0,60	0,70	0,80	0,90	0,95	0,975	0,990	0,995	0,999	0,9995	
1	0,325	0,727	1,376	3,078	6,314	12,71	31,82	63,66	318,3	636,6	1
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925	22,33	31,60	2
3	0,277	0,584	0,978	1,638	2,353	3,182	4,541	5,841	10,21	12,92	3
4	0,271	0,569	0,941	1,533	2,132	2,776	3,747	4,604	7,173	8,610	4
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	4,032	5,893	6,869	5
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707	5,208	5,959	6
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499	4,785	5,408	7
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355	4,501	5,041	8
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250	4,297	4,781	9
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169	4,144	4,587	10
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106	4,025	4,437	11
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055	3,930	4,318	12
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012	3,852	4,221	13
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	2,977	3,787	4,140	14
15	0,258	0,536	0,866	1,341	1,753	2,131	2,602	2,947	3,733	4,073	15
16	0,258	0,535	0,865	1,337	1,746	2,120	2,583	2,921	3,686	4,015	16
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898	3,646	3,965	17
18	0,257	0,534	0,862	1,330	1,734	2,101	2,552	2,878	3,610	3,922	18
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861	3,579	3,883	19
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845	3,552	3,850	20
21	0,257	0,532	0,859	1,323	1,721	2,080	2,518	2,831	3,527	3,819	21
22	0,256	0,532	0,858	1,321	1,717	2,074	2,508	2,819	3,505	3,792	22
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807	3,485	3,768	23
24	0,256	0,531	0,857	1,318	1,711	2,064	2,492	2,797	3,467	3,745	24
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787	3,450	3,725	25
26	0,256	0,531	0,856	1,315	1,706	2,056	2,479	2,779	3,435	3,707	26
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771	3,421	3,690	27
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763	3,408	3,674	28
29	0,256	0,530	0,854	1,311	1,699	2,045	2,462	2,756	3,396	3,659	29
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750	3,385	3,646	30
32	0,255	0,530	0,853	1,309	1,694	2,037	2,449	2,738	3,365	3,622	32
34	0,255	0,529	0,852	1,307	1,691	2,032	2,441	2,728	3,348	3,601	34
36	0,255	0,529	0,852	1,306	1,688	2,028	2,434	2,719	3,333	3,582	36
38	0,255	0,529	0,851	1,304	1,686	2,024	2,429	2,712	3,319	3,566	38
40	0,255	0,529	0,851	1,303	1,684	2,021	2,423	2,704	3,307	3,551	40
50	0,255	0,528	0,849	1,299	1,676	2,009	2,403	2,678	3,261	3,496	50
60	0,254	0,527	0,848	1,296	1,671	2,000	2,390	2,660	3,232	3,460	60
70	0,254	0,527	0,847	1,294	1,667	1,994	2,381	2,648	3,211	3,435	70
80	0,254	0,526	0,846	1,292	1,664	1,990	2,374	2,639	3,195	3,416	80
90	0,254	0,526	0,846	1,291	1,662	1,987	2,368	2,632	3,183	3,402	90
100	0,254	0,526	0,845	1,290	1,660	1,984	2,364	2,626	3,174	3,390	100
200	0,254	0,525	0,843	1,286	1,653	1,972	2,345	2,601	3,131	3,340	200
500	0,253	0,525	0,842	1,283	1,648	1,965	2,334	2,586	3,107	3,310	500
∞	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	3,090	3,291	∞
	0,60	0,70	0,80	0,90	0,95	0,975	0,990	0,995	0,999	0,9995	

Table des quantiles d'ordre 0,95 des lois de Fisher-Snédécor $F(m_1,m_2)$

Nombre de degrés de liberté du dénominateur: m_2	Nombre de degrés de liberté du numérateur: m_1																									
	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	22	24	26	28	30	40	60	80	100	∞	
1	161	200	216	225	230	234	237	239	241	242	244	245	246	247	248	249	249	249	250	250	251	252	253	253	254	1
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	2
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,71	8,69	8,67	8,66	8,65	8,64	8,63	8,62	8,62	8,59	8,57	8,56	8,55	8,53	3
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,87	5,84	5,82	5,80	5,79	5,77	5,76	5,75	5,75	5,72	5,69	5,67	5,66	5,63	4
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,64	4,60	4,58	4,56	4,54	4,53	4,52	4,50	4,50	4,46	4,43	4,41	4,41	4,37	5
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,96	3,92	3,90	3,87	3,86	3,84	3,83	3,82	3,81	3,77	3,74	3,72	3,71	3,67	6
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,53	3,49	3,47	3,44	3,43	3,41	3,40	3,39	3,38	3,34	3,30	3,29	3,27	3,23	7
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,24	3,20	3,17	3,15	3,13	3,12	3,10	3,09	3,08	3,04	3,01	2,99	2,97	2,93	8
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,03	2,99	2,96	2,94	2,92	2,90	2,89	2,87	2,86	2,83	2,79	2,77	2,76	2,71	9
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,86	2,83	2,80	2,77	2,75	2,74	2,72	2,71	2,70	2,66	2,62	2,60	2,59	2,54	10
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,74	2,70	2,67	2,65	2,63	2,61	2,59	2,58	2,57	2,53	2,49	2,47	2,46	2,40	11
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,64	2,60	2,57	2,54	2,52	2,51	2,49	2,48	2,47	2,43	2,38	2,36	2,35	2,30	12
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,55	2,51	2,48	2,46	2,44	2,42	2,41	2,39	2,38	2,34	2,30	2,27	2,26	2,21	13
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,48	2,44	2,41	2,39	2,37	2,35	2,33	2,32	2,31	2,27	2,22	2,20	2,19	2,13	14
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,42	2,38	2,35	2,33	2,31	2,29	2,27	2,26	2,25	2,20	2,16	2,14	2,12	2,07	15
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,37	2,33	2,30	2,28	2,25	2,24	2,22	2,21	2,19	2,15	2,11	2,08	2,07	2,01	16
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,33	2,29	2,26	2,23	2,21	2,19	2,17	2,16	2,15	2,10	2,06	2,03	2,02	1,96	17
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,29	2,25	2,22	2,19	2,17	2,15	2,13	2,12	2,11	2,06	2,02	1,99	1,98	1,92	18
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,26	2,21	2,18	2,16	2,13	2,11	2,10	2,08	2,07	2,03	1,98	1,96	1,94	1,88	19
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,22	2,18	2,15	2,12	2,10	2,08	2,07	2,05	2,04	1,99	1,95	1,92	1,91	1,84	20
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,20	2,16	2,12	2,10	2,07	2,05	2,04	2,02	2,01	1,96	1,92	1,89	1,88	1,81	21
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,17	2,13	2,10	2,07	2,05	2,03	2,01	2,00	1,98	1,94	1,89	1,86	1,85	1,78	22
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,15	2,11	2,08	2,05	2,02	2,01	1,99	1,97	1,96	1,91	1,86	1,84	1,82	1,76	23
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,13	2,09	2,05	2,03	2,00	1,98	1,97	1,95	1,94	1,89	1,84	1,82	1,80	1,73	24
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,11	2,07	2,04	2,01	1,98	1,96	1,95	1,93	1,92	1,87	1,82	1,80	1,78	1,71	25
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,09	2,05	2,02	1,99	1,97	1,95	1,93	1,91	1,90	1,85	1,80	1,78	1,76	1,69	26
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,08	2,04	2,00	1,97	1,95	1,93	1,91	1,90	1,88	1,84	1,79	1,76	1,74	1,67	27
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,06	2,02	1,99	1,96	1,93	1,91	1,90	1,88	1,87	1,82	1,77	1,74	1,73	1,65	28
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,05	2,01	1,97	1,94	1,92	1,90	1,88	1,87	1,85	1,81	1,75	1,73	1,71	1,64	29
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,04	1,99	1,96	1,93	1,91	1,89	1,87	1,85	1,84	1,79	1,74	1,71	1,70	1,62	30
32	4,15	3,29	2,90	2,67	2,51	2,40	2,31	2,24	2,19	2,14	2,07	2,01	1,97	1,94	1,91	1,88	1,86	1,85	1,83	1,82	1,77	1,71	1,69	1,67	1,59	32
34	4,13	3,28	2,88	2,65	2,49	2,38	2,29	2,23	2,17	2,12	2,05	1,99	1,95	1,92	1,89	1,86	1,84	1,82	1,81	1,80	1,75	1,69	1,66	1,65	1,57	34
36	4,11	3,26	2,87	2,63	2,48	2,36	2,28	2,21	2,15	2,11	2,03	1,98	1,93	1,90	1,87	1,85	1,82	1,81	1,79	1,78	1,73	1,67	1,64	1,62	1,55	36
38	4,10	3,24	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	2,02	1,96	1,92	1,88	1,85	1,83	1,81	1,79	1,77	1,76	1,71	1,65	1,62	1,61	1,53	38
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,95	1,90	1,87	1,84	1,81	1,79	1,77	1,76	1,74	1,69	1,64	1,61	1,59	1,51	40
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,95	1,89	1,85	1,81	1,78	1,76	1,74	1,72	1,70	1,69	1,63	1,58	1,54	1,52	1,44	50
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,86	1,82	1,78	1,75	1,72	1,70	1,68	1,66	1,65	1,59	1,53	1,50	1,48	1,39	60
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,89	1,84	1,79	1,75	1,72	1,70	1,67	1,65	1,64	1,62	1,57	1,50	1,47	1,45	1,35	70
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,88	1,82	1,77	1,73	1,70	1,68	1,65	1,63	1,62	1,60	1,54	1,48	1,45	1,43	1,32	80
90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99	1,94	1,86	1,80	1,76	1,72	1,69	1,66	1,64	1,62	1,60	1,59	1,53	1,46	1,43	1,41	1,30	90
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,85	1,79	1,75	1,71	1,68	1,65	1,63	1,61	1,59	1,57	1,52	1,45	1,41	1,39	1,28	100
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,69	1,64	1,60	1,57	1,54	1,52	1,50	1,48	1,46	1,39	1,32	1,27	1,24	1,00	∞
	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	22	24	26	28	30	40	60	80	100	∞	