TD 2 : Estimation

Exercice 1 ***Manipulation de lois normales, et lecture de table*** *Soit* $X \sim \mathcal{N}(0,1)$,

1. *Lire dans la table* $p_1 = \mathbb{P}(X \leq 1.24)$ *et* q_1 *tel que* $\mathbb{P}(X \leq q_1) = 0.75175$.
2. *Soit* $t > 0$, *que vaut* $\mathbb{P}(X = t)$ *? Exprimer* $\mathbb{P}(X \geq t)$ *en fonction de* $\mathbb{P}(X \leq t)$.
3. *Soit* $t < 0$ *exprimer* $\mathbb{P}(X \leq -t)$ *en fonction de* $\mathbb{P}(X \leq t)$.
4. *Lire dans la table* $p_2 = \mathbb{P}(X \leq -1.04)$, *et* q_2 *tel que* $\mathbb{P}(X \leq q_2) = 0.45224$.
5. *Soit* $t > 0$, *Exprimer* $\mathbb{P}(|X| < t)$ *en fonction de* $\mathbb{P}(X \leq t)$.
6. *Lire dans la table* $p_3 = \mathbb{P}(|X| \leq 1.96)$ *et* q_3 *tel que* $\mathbb{P}(|X| \geq q_3) = 0.101$.
7. *Quelle est la loi de* $Z = 2X + 1$ *?*
8. *Lire dans la table* $p_4 = \mathbb{P}(Z \leq 3.56)$ *et* q_4 *tel que* $\mathbb{P}(Z > q_4) = 0.04006$.

Soit $X_1 \sim \mathcal{N}(1,9)$ *et* $X_2 \sim \mathcal{N}(-1,4)$, *indépendantes.*

1. *Quelle est la loi de* $Y = -2(X_1 - 2X_2)$ *?*
2. *Centrer et réduire* Y.
3. *En utilisant la centrée réduite de* Y, *calculer* $\mathbb{P}(Y > 13.6)$
4. *En utilisant la centrée réduite de* Y, *trouver* q *tel que* $\mathbb{P}(|Y + 6| > q) = 0.89656$.

Exercice 2 ***Les premiers estimateurs***

Soient X_1, X_2 *deux variables i.i.d., d'espérance* m *inconnue, et de variance* σ^2 *(que l'on suppose connue). On pose* $\hat{M}_1 = m, \hat{M}_2 = 1$, $\hat{M}_3 = 2X_1 + X_2, \hat{M}_4 = \frac{2X_1+X_2}{3}, \hat{M}_5 = \frac{X_1+X_2}{2}$.

1. *Parmi les* $\hat{M}_i$, *lesquels sont des statistiques ?*
2. *Parmi les* $\hat{M}_i$, *lesquels sont des estimateurs sans biais de* m *?*
3. *Parmi les estimateurs sans biais, lequel à la plus petite variance ?*
4. *Quelle est la variance de* $\hat{M}_2$ *? Commenter.*

Exercice 3 ***Les premiers estimateurs***

Soient X_1, X_2 *deux variables indépendantes, d'espérance* m *inconnue de variances respectives* σ_1^2 *et* σ_2^2, *connues. On cherche à définir un estimateur de* m *de la forme* $\hat{M} = aX_1 + bX_2$.

1. *Trouver une condition sur* a *et* b *pour que* $\hat{M}$ *soit sans biais.*
2. *Calculer la variance de* $\hat{M}$ *en fonction de* a, b, σ_1 *et* σ_2.
3. *En utilisant la condition d'estimateur non biaisé (par exemple, en remplaçant* b *en fonction de* a*) dans l'expression de la variance, donner l'estimateur sans biais de variance minimale.*

Exercice 4 ***Principe d'un intervalle de confiance et idée d'un test statistique***

Un déodorant affiche une durée de protection moyenne de $6h$. *Un institut prend* 16 *mesures (que l'on considère comme fiables) de durées de protection, et on admet que l'on peut modéliser ces durées comme des réalisations i.i.d. de lois normales d'espérance* m *inconnue, et d'écart-type* 1 *(en heures). Soit* $X_1, \cdots, X_{16}$ *les mesures prises.*

1. *On pose* $\bar{X} = \frac{1}{16}\sum_{i=1}^{16} X_i$. *Calculer* $\mathbb{E}[\bar{X}]$
2. $\bar{X}$ *est-il un estimateur sans biais de* m *?*
3. *Sous les hypothèses de l'énoncé, quelle est la loi de* $\bar{X}$ *(en fonction de* m, *inconnu) ?*
4. *Quelle est la loi de la variable* $T = 4(\bar{X} - m)$ *?*
5. *À l'aide des tables de quantiles, trouver* a *tel que* $\mathbb{P}(|T| > a) = 0.1$.
6. *En déduire un intervalle aléaoire de la forme* $I = [A_1(X_1, ..., X_{16}), A_2(X_1, ..., X_{16})]$ *tel que* $\mathbb{P}(m \in I) = 0.9$

Exercice 5 ***Sondage***

Du contexte pour la culture :

Pour les sondages politiques, il est d'usage d'interroger 1000 *personnes, la plupart du temps avec la méthode dite "des quotas" qui consiste à essayer de tirer un échantillon "représentatif" de la population (au sens où l'échantillonnage cherche à reproduire dans les mêmes proportions que dans la population globale les classes d'âges, origines (régions, taille de la commune...), les CSP, les sexes...).*

Cette méthode n'a de fondement mathématique que si l'on exclut aucune variable qui pourrait influencer le résultat que l'on cherche à prédire (vainqueur d'une élection par exemple). En effet, cette méthode permet de s'affranchir de l'hypothèse d'aléa "On a échantillonné uniformément dans la population" (chose difficile à faire en pratique), mais enlève la validité mathématique des résultats, sans hypothèse forte sur le choix des variables dites de "stratification".

En pratique, les sondeurs calculent leur marge d'erreur enconsidérant qu'ils feront au moins aussi bien que la méthode d'échantillonage aléatoire.

Calcul de l'erreur d'un sondage avec une méthode d'échantillonage aléatoire:

On suppose que l'on interroge 1000 *personnes tirées au hasard, uniformément dans la population (et indépendamment, avec remise), qu'elles acceptent toutes de répondre, qu'elles ne mentent pas, et que parmi elles* k *disent qu'elles vont voter pour le candidat A, et* $1000-k$ *pour le candidat B.*

1. *Par quelle loi peut on modéliser le choix* X_1 *du premier individu (pris* 1 *pour le candidat A, et* 0 *pour le candidat B pour simplifier) ? On notera* p *le paramètre de cette loi.*
2. *On note* $n = 1000$, *et* $X_1, \cdots, X_n$ *les choix des individus . Quelle est la loi de* $S_n = \sum_{i=1}^n X_i$ *?*
3. *Que vallent, en fonction de* p, $\mathbb{P}(S_n = 0)$ *et* $\mathbb{P}(S_n = 1)$ *?*
4. *On pose* $\bar{X}_n = \frac{S_n}{n}$. *Vers quoi converge en loi* $Z_n = \sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}}$ *? On désigne par* Z *une v.a. de cette loi.*
5. *Trouver* a *tel que* $\mathbb{P}(|Z| > a) = 0.01$.
6. *En déduire* $b(n, p)$ *tel que* $\mathbb{P}\Big(|\bar{X}_n - p| > b(n, p)\Big) \to 0.01$.
7. *En utilisant* $p(1-p) \leq 0.25, \forall p \in [0, 1]$, *en déduire* c *tel que* $\limsup \mathbb{P}\Big(|\bar{X}_n - p| > c\Big) \leq 0.01$. *Commenter*

Exercice 6 ***Avec la loi exponentielle***

Un fabriquant d'ampoules explique que la durée de vie de ses ampoules suit une loi exponentielle.

On a fait l'expérience en mesurant la durée d'utilisation de $n = 20$ *ampoules, qu'on a modélisées par des réalisations* $y_1, \cdots, y_n$ *de v.a.* $Y_1, \cdots, Y_n$ *de loi exponentielle i.i.d. de paramètre* λ *(inconnu).*

1. *En fonction de* λ, *donner la durée de vie moyenne d'une ampoule.*
2. *Soit* $s > 0$ *un temps donné. Calculer, en fonction de* λ, *la probabilité qu'une ampoule dure plus de* s, *en déduire la probabilité que la durée de vie d'une ampoule soit supérieure à* $\frac{2}{\lambda}$.
3. *Calculer le quantile*[1] *d'ordre* p *de de la variable aléatoire* Y_1, *en déduire la médiane de* Y_1.
4. *Soit* $s, t > 0$ *Calculer, en fonction de* λ, *la probabilité qu'une ampoule soit encore fonctionnelle au temps* $s + t$ *sachant qu'elle l'était au temps* t.
5. *La seule information dont vous disposez, c'est que sur les* 20 *ampoules,* 4 *ont duré plus de* 3465 *heures. Le fabriquant affiche une durée moyenne de* 5000 *heures. En posant* $B_i = \{Y_i > 3465\}, i = 1 \cdots 20$, *quelle serait la probabilité, si le fabriquant disait la vérité sur la durée moyenne, que seulement* 4 *évènements ou moins soient survenus parmis les* B_i.

Exercice 7 ***Simple Intervalle de Confiance***

On cherche à contrôler la température dans un processus industriel. On prend simultanément 25 *mesures de température, que l'on modélise comme des lois normales i.i.d. d'espérance* m *inconnue, et d'écart type* $\sigma = 0.2°C$ *connu. On déclenchera la régulation dès que la température de* $18°C$ *ne tombe pas dans l'intervalle de confiance.*

1. *Construire un IC bilatéral sur* m *à* 99%.
2. *La moyenne des mesures prises est de* $\bar{x} = 18.1°C$. *Donner l'intervalle de confiance ponctuel obtenu.*
3. *On prend maintenant des mesures indépendantes chaque minute. Même si la température* m *ne bouge pas durant tout ce temps, que va-t-il se passer ?*

[1] i.e. q tel que $\mathbb{P}(Y_1 \leq q) = p$

Table de la loi normale centrée réduite

La table donne en fonction de z (qui se lit en ligne pour la partie entière et première décimale, et en colonne pour la seconde), la fonction de répartition $\mathbb{P}(Z \leq z)$.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.975	0.97558	0.97615	0.9767
2	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.983	0.98341	0.98382	0.98422	0.98461	0.985	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999
3.1	0.99903	0.99906	0.9991	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.9994	0.99942	0.99944	0.99946	0.99948	0.9995
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.9996	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.9997	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.9998	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.9999	0.9999	0.9999	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998