Chapitre 8

Le theorème limite central

Avec la loi des grands nombres nous avons vu le lien qui existe entre la notion abstraite de probabilité d'un évènement, et la fréquence de l'occurence de cet évenement dans une suite de réalisations indépendantes et identiques d'une expérience pouvant provoquer cet évènement (Théorème de Bernoulli). Plus génralement, la loi des grands nombre montre que la moyenne $M_n := \frac{1}{n}\sum_{i=1}^n X_i$ d'une suite de v.a. X_i i.i.d. (ayant espérance et variance) tends vers un nombre : l'espèrance commune $\mu := \mathbb{E}(X_i)$ de ces v.a. ; bien entendu, M_n n'est que de variance petite et reste donc aléatoire : si l'on veut observer la variabilité de M_n (qui est proche de μ) il est nécéssaire d'agrandir l'écart $M_n - \mu$ en le multipliant par une grandeur $\lambda(n)$ suffisament grande avec n. Le choix de $\lambda(n)$ peut se faire simplement de manière à retrouver dans la v.a. "amplifiée" $Z_n := \lambda(n)(M_n - \mu)$ soit de variance 1. Calculons : nous voulons

$$\begin{aligned}
1 &= \operatorname{Var}(Z_n) = \operatorname{Var}(\lambda(n)(M_n-\mu)) = \lambda^2(n)\operatorname{Var}((M_n-\mu)) = \lambda^2(n)\operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right)\\
&= \frac{\lambda^2(n)}{n^2}\operatorname{Var}\left(\sum_{i=1}^n X_i\right) \text{ par indépendance des } X_i\\
&= \frac{\lambda^2(n)}{n^2}\operatorname{Var}\left(\sum_{i=1}^n X_i\right) = \frac{\lambda^2(n)}{n^2}\sum_{i=1}^n \operatorname{Var}(X_i) = \frac{\lambda^2(n)}{n^2} n\operatorname{Var}(X_i) = \frac{\lambda^2(n)}{n}\sigma^2.
\end{aligned}$$

Il suffit donc de choisir $\boxed{\lambda(n) := \frac{\sqrt{n}}{\sigma}}$ pour que Z_n soit de variance 1. Nous avons alors

$$Z_n := \lambda(n)(M_n-\mu) = \frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu\right) = \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^n X_i - n\mu\right)$$

qui est visiblement d'espérance nulle. Et là se produit une petite merveille : pour n grand, cette "remise à l'échelle" de $M_n - \mu$ produit une v.a. Z_n *dont la loi est proche de la loi normale* (centrée et réduite, puisque tel est le cas pour la loi de tous les Z_n), et ceci sans autre hypothèse : ceci explique pourquoi cette loi surgit si souvent en statistique. Cette merveille porte le nom de théorème limite central[1] (central limit theorem, en anglais)

8.1 Le théorème

Théorème 8.1 (théorème limite central) *Avec les notations et les hypothèses ci-dessus,* $Z_n := \frac{1}{\sigma\sqrt{n}}(X_1 + \ldots + X_n - n\mu)$ *tend en loi vers une v.a.* Z *de loi* $\mathcal{N}(0,1)$

Preuve : L'idée de la preuve que nous allons donner est simple : nous allons montrer que la fonction génératrice des moments M_{Z_n} des Z_n tends vers la fonction génératrice des moments M_Z de Z. Le théorème 7.7 permettra alors de conclure, puis qu'il assure que ceci implique la convergence en loi des Z_n vers Z. Nous utiliserons le lemme général suivant, dont nous laissons la preuve, élémentaire, en exercice :

[1]Nous adoptons ici la traduction *thèorème limite central* de "central limit theorem" donnée par Jean Jacod et Philip Protter dans *L'essentiel en théorie des probabilités* (Cassini, Paris 2003) et qui est la traduction de leur *Probability Essentials* (Springer, Berlin et al. 2000). La traduction théorème central limit est souvent adoptée et a donné l'acronyme "TCL" pour désigner ce théorème ; celle-ci est malheureuse, car il s'agit bien ici d'un théorème limite qui a un rôle central dans la théorie : le nom original est donc à comprendre par (central (limit theorem)) et non ((central limit) theorem) ; gâre à la non-associativité en linguistique !

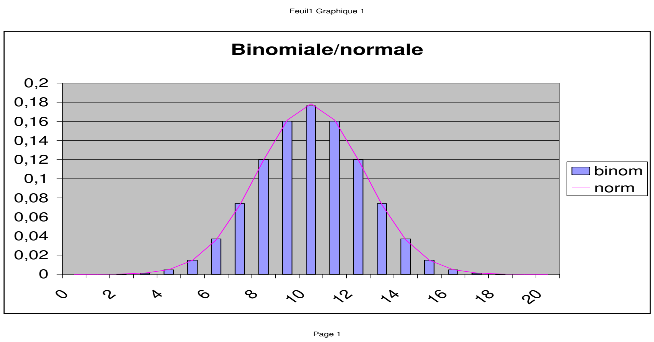

FIG. 8.1 – Le théorème limite central a été découvert par de Moivre et démontré par Laplace dans le cas d'une suite de v.a. de Bernoulli i.i.d. ; l'approximation d'une loi binômilale par une loi normale qui en résulte porte dès-lors le nom de théorème de de Moivre-Laplace. Ci dessus un histogramme de la loi binômiale $\mathcal{B}(20, 0.5)$ et de la densité de la loi normale $\mathcal{N}(\mu, \sigma)$ de même espérance $\mu = np = 10$ et même écart-type $\sigma = \sqrt{np(1-p)} = \sqrt{5}$.

Lemme 8.2 *Pour tout v.a. X pour laquelle M_X est définie, et tout $a \in \mathbb{R}$, on a* $\boxed{M_{aX}(t) = M_X(at)}$.

Rappelons que

$$Z_n := \frac{1}{\sigma\sqrt{n}}(X_1 + \ldots + X_n - n\mu) = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu),$$

et comme les v.a. $X_i - \mu$ sont indépendantes, par la proposition 7.6 on a

$$M_{Z_n}(t) = M_{\frac{1}{\sigma\sqrt{n}}\sum_{i=1}^n (X_i-\mu)}(t) = M_{\sum_{i=1}^n (X_i-\mu)}(\frac{t}{\sigma\sqrt{n}}) = \prod_{i=1}^{n} M_{X_i-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right) = \left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n,$$

où l'on a posé $\varphi(\tau_n) := M_{X_i-\mu}(\tau_n)$ (qui sont toutes égales puisque les v.a. ont même loi), et $\tau_n = \frac{t}{\sigma\sqrt{n}}$; observons que $\lim_{n\to+\infty}\tau_n = \lim_{n\to+\infty}\frac{t}{\sigma\sqrt{n}} = 0$ pour tout t. Pour calculer la limite des $M_{Z_n}(t)$, nous alons utiliser un développement de Taylor à l'ordre 2 de φ ; pour cela, il nous faut connaître $\varphi(0)$, $\varphi'(0)$, et $\varphi''(0)$, que nous calculons en utilisant la formule (7.2). On a

$$\begin{aligned} \varphi(0) &= \mathbb{E}((X_i-\mu)^0) = \mathbb{E}(1) = 1, \\ \varphi'(0) &= \mathbb{E}(X_i-\mu) = \mathbb{E}(X_i) - \mu = 0, \\ \varphi''(0) &= \mathbb{E}((X_i-\mu)^2) = \mathrm{Var}\,(X_i) = \sigma^2, \end{aligned}$$

d'où finalement $\varphi(\tau) = \varphi(0) + \tau\varphi'(0) + \frac{\tau^2}{2}(\varphi''(0) + \varepsilon(\tau))$, où $\lim_{\tau\to 0}\varepsilon(\tau) = 0$. Nous allons également utiliser un développement de Taylor de $\ln(1+u)$; nous s'avons qu'il existe une fonction η telle que $\ln(1+u) = u(1+\eta(u))$, avec $\lim_{u\to 0}\eta(u) = 0$. Nous avons donc

$$\begin{aligned} M_{Z_n}(t) &= \left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n = \left(1 + \frac{t^2}{2\sigma^2 n}(\sigma^2+\varepsilon_n)\right)^n \text{ , où } \varepsilon_n := \varepsilon\left(\tfrac{t}{\sigma\sqrt{n}}\right) \\ &= \exp\left(n\ln\left(1 + \frac{t^2}{2\sigma^2 n}(\sigma^2+\varepsilon_n)\right)\right) \\ &= \exp\left(n\left(\frac{t^2}{2\sigma^2 n}(\sigma^2+\varepsilon_n)\right)(1+\eta_n)\right) \text{ , avec } \eta_n = \eta\left(\tfrac{t^2}{2\sigma^2 n}(\sigma^2+\varepsilon_n)\right) \\ &= \exp\left(\left(\frac{t^2}{2\sigma^2}(\sigma^2+\varepsilon_n)\right)(1+\eta_n)\right). \end{aligned}$$

Donc

$$\lim_{n\to+\infty} M_{Z_n}(t) = \lim_{n\to+\infty}\exp\left(\left(\frac{t^2}{2\sigma^2}(\sigma^2+\varepsilon_n)\right)(1+\eta_n)\right) = e^{\frac{t^2}{2\sigma^2}\sigma^2} = e^{\frac{t^2}{2}} = M_Z(t).$$

On conclut, comme annoncé, en appliquant le théoreme 7.7.

□

8.2 Pratique du théorème limite central

Exemple typique : Une compagnie aérienne donne des réservations sur le vol d'un appareil de 400 places. La probabilité qu'un passager ayant réservé pour ce vol ne se présente pas est de $0.08 = 8\%$. Si la compagnie accorde 420 réservations sur ce vol, quel est le risque de "surbooking" (c'est-à-dire qu'il se présente plus de passagers que les 400 qui pourront embarquer) ?

Résolution par approximation normale : Soit X_i la v.a. de Bernoulli modélisant la présence ($X_i = 1$) ou non ($X_i = 0$) du i-ème passager réservé, $i = 1..n$, avec $n = 420$; par hypothèse $X_i \rightsquigarrow \mathcal{B}(1, p)$, avec $p = 1 - 8\% = 0.92\%$, et on suppose (implicitement ...) que les X_i sont indépendants. On a donc $\mu = \mathbb{E}X_i = p$, et $\sigma = \sqrt{\mathrm{Var}\,(X_i)} = \sqrt{p(1-p)}$.

Soit $X := X_1 + .. + X_{420}$ le nombre (aléatoire) de passagers réservés se présentant effectivement à l'enregistrement. Sous nos hypothèses $\mathbb{E}X = np = 420 \cdot 0.92 = 384.4$ et $\mathrm{Var}\,(X) = np(1-p) = 420 \cdot 0.92 \cdot 0.08 = 30.912$. Soit

$$\boxed{Z_n = \frac{X - \mathbb{E}X}{\sqrt{\mathrm{Var}\,(X)}}} \left(= \frac{X_1 + \ldots + X_{420} - n\mu}{\sigma\sqrt{n}} \right)$$

la v.a. considérée dans le théorème limite central, qui n'est autre que X centrée et réduite. L'*application du theorème consiste à assimiler* Z_{420} à $Z \rightsquigarrow \mathcal{N}(0, 1)$.

L'évènement dont nous recherchons la probabilité est

$$E := \{X \leq 400\} = \left\{ Z_{420} \leq \frac{400 - \mathbb{E}X}{\sqrt{\mathrm{Var}\,(X)}} \right\} \simeq \left\{ Z \leq \frac{400 - \mathbb{E}X}{\sqrt{\mathrm{Var}\,(X)}} \right\} = \{Z \lessapprox 2.45\},$$

puisque $\frac{400-\mathbb{E}X}{\sqrt{\mathrm{Var}\,(X)}} = \frac{400-np}{\sqrt{np(1-p)}} = \frac{400-384.4}{\sqrt{30.912\ldots}} = 2.446\ldots$.

Comme $Z \rightsquigarrow \mathcal{N}(0, 1)$, on recherche la valeur $2.45 = 2.40 + 0.05 = u_1 + u_2$ dans la table ci-dessous, et on trouve $\mathbb{P}(E) = 0.992857\ldots$. Il y a donc moins de 1% de risque qu'il se présente plus de 400 passagers à l'enregistrement[2].

[2]Sur la loi européenne sur la surréservation, voir par exemple :
http ://www.europe.gouv.fr/europe_7/europe_au_quotidien_25/surreservation_surbooking_avion_136.html

Voici quelques valeurs de la loi normale centrée réduite $\mathcal{N}(0,1)$ calculée au moyen d'Excel :

Loi Normale N(0,1)
Probabilité que X soit inférieure à u1+u2

u1\u2	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000000	0,5039894	0,5079784	0,5119665	0,5159535	0,5199389	0,5239223	0,5279032	0,5318814	0,5358565
0,1	0,5398279	0,5437954	0,5477585	0,5517168	0,5556700	0,5596177	0,5635595	0,5674949	0,5714237	0,5753454
0,2	0,5792597	0,5831661	0,5870644	0,5909541	0,5948348	0,5987063	0,6025681	0,6064198	0,6102612	0,6140918
0,3	0,6179114	0,6217195	0,6255158	0,6293000	0,6330717	0,6368306	0,6405764	0,6443087	0,6480272	0,6517317
0,4	0,6554217	0,6590970	0,6627572	0,6664021	0,6700314	0,6736448	0,6772419	0,6808225	0,6843863	0,6879331
0,5	0,6914625	0,6949743	0,6984682	0,7019441	0,7054015	0,7088403	0,7122603	0,7156612	0,7190427	0,7224047
0,6	0,7257469	0,7290692	0,7323712	0,7356528	0,7389138	0,7421540	0,7453732	0,7485712	0,7517478	0,7549030
0,7	0,7580364	0,7611480	0,7642376	0,7673050	0,7703501	0,7733727	0,7763728	0,7793501	0,7823046	0,7852362
0,8	0,7881447	0,7910300	0,7938920	0,7967307	0,7995459	0,8023375	0,8051055	0,8078498	0,8105704	0,8132671
0,9	0,8159399	0,8185888	0,8212136	0,8238145	0,8263912	0,8289439	0,8314724	0,8339768	0,8364569	0,8389129
1,0	0,8413447	0,8437523	0,8461358	0,8484950	0,8508300	0,8531409	0,8554277	0,8576903	0,8599289	0,8621434
1,1	0,8643339	0,8665004	0,8686431	0,8707618	0,8728568	0,8749280	0,8769755	0,8789995	0,8809998	0,8829767
1,2	0,8849303	0,8868605	0,8887675	0,8906514	0,8925122	0,8943502	0,8961653	0,8979576	0,8997274	0,9014746
1,3	0,9031995	0,9049020	0,9065824	0,9082408	0,9098773	0,9114919	0,9130850	0,9146565	0,9162066	0,9177355
1,4	0,9192433	0,9207301	0,9221961	0,9236414	0,9250663	0,9264707	0,9278549	0,9292191	0,9305633	0,9318879
1,5	0,9331928	0,9344783	0,9357445	0,9369916	0,9382198	0,9394292	0,9406200	0,9417924	0,9429466	0,9440826
1,6	0,9452007	0,9463011	0,9473839	0,9484493	0,9494974	0,9505285	0,9515428	0,9525403	0,9535214	0,9544861
1,7	0,9554346	0,9563671	0,9572838	0,9581849	0,9590705	0,9599409	0,9607961	0,9616365	0,9624621	0,9632731
1,8	0,9640697	0,9648522	0,9656206	0,9663751	0,9671159	0,9678433	0,9685573	0,9692582	0,9699460	0,9706211
1,9	0,9712835	0,9719335	0,9725711	0,9731967	0,9738102	0,9744120	0,9750022	0,9755809	0,9761483	0,9767046
2,0	0,9772499	0,9777845	0,9783084	0,9788218	0,9793249	0,9798179	0,9803008	0,9807739	0,9812373	0,9816912
2,1	0,9821356	0,9825709	0,9829970	0,9834143	0,9838227	0,9842224	0,9846137	0,9849966	0,9853713	0,9857379
2,2	0,9860966	0,9864475	0,9867907	0,9871263	0,9874546	0,9877756	0,9880894	0,9883962	0,9886962	0,9889894
2,3	0,9892759	0,9895559	0,9898296	0,9900969	0,9903582	0,9906133	0,9908625	0,9911060	0,9913437	0,9915758
2,4	0,9918025	0,9920237	0,9922397	0,9924506	0,9926564	0,9928572	0,9930531	0,9932443	0,9934309	0,9936128
2,5	0,9937903	0,9939634	0,9941322	0,9942969	0,9944574	0,9946138	0,9947664	0,9949150	0,9950600	0,9952012
2,6	0,9953388	0,9954729	0,9956035	0,9957307	0,9958547	0,9959754	0,9960929	0,9962074	0,9963188	0,9964274
2,7	0,9965330	0,9966358	0,9967359	0,9968332	0,9969280	0,9970202	0,9971099	0,9971971	0,9972820	0,9973645
2,8	0,9974448	0,9975229	0,9975988	0,9976725	0,9977443	0,9978140	0,9978817	0,9979476	0,9980116	0,9980737
2,9	0,9981341	0,9981928	0,9982498	0,9983051	0,9983589	0,9984111	0,9984617	0,9985109	0,9985587	0,9986050
3,0	0,9986500	0,9986937	0,9987361	0,9987772	0,9988170	0,9988557	0,9988932	0,9989296	0,9989649	0,9989991
3,1	0,9990323	0,9990645	0,9990957	0,9991259	0,9991552	0,9991836	0,9992111	0,9992377	0,9992636	0,9992886
3,2	0,9993128	0,9993363	0,9993590	0,9993810	0,9994023	0,9994229	0,9994429	0,9994622	0,9994809	0,9994990
3,3	0,9995165	0,9995335	0,9995499	0,9995657	0,9995811	0,9995959	0,9996102	0,9996241	0,9996375	0,9996505
3,4	0,9996630	0,9996751	0,9996868	0,9996982	0,9997091	0,9997197	0,9997299	0,9997397	0,9997492	0,9997584
3,5	0,9997673	0,9997759	0,9997842	0,9997922	0,9997999	0,9998073	0,9998145	0,9998215	0,9998282	0,9998346
3,6	0,9998409	0,9998469	0,9998527	0,9998583	0,9998636	0,9998688	0,9998739	0,9998787	0,9998834	0,9998878
3,7	0,9998922	0,9998963	0,9999004	0,9999042	0,9999080	0,9999116	0,9999150	0,9999184	0,9999216	0,9999247
3,8	0,9999276	0,9999305	0,9999333	0,9999359	0,9999385	0,9999409	0,9999433	0,9999456	0,9999478	0,9999499
3,9	0,9999519	0,9999538	0,9999557	0,9999575	0,9999592	0,9999609	0,9999625	0,9999640	0,9999655	0,9999669
4,0	0,9999683	0,9999696	0,9999709	0,9999721	0,9999733	0,9999744	0,9999755	0,9999765	0,9999775	0,9999784
4,1	0,9999793	0,9999802	0,9999810	0,9999819	0,9999826	0,9999834	0,9999841	0,9999848	0,9999854	0,9999860
4,2	0,9999866	0,9999872	0,9999878	0,9999883	0,9999888	0,9999893	0,9999898	0,9999902	0,9999906	0,9999911
4,3	0,9999915	0,9999918	0,9999922	0,9999925	0,9999929	0,9999932	0,9999935	0,9999938	0,9999941	0,9999943
4,4	0,9999946	0,9999948	0,9999951	0,9999953	0,9999955	0,9999957	0,9999959	0,9999961	0,9999963	0,9999964
4,5	0,9999966	0,9999968	0,9999969	0,9999970	0,9999972	0,9999973	0,9999974	0,9999976	0,9999977	0,9999978
4,6	0,9999979	0,9999980	0,9999981	0,9999982	0,9999983	0,9999983	0,9999984	0,9999985	0,9999986	0,9999986
4,7	0,9999987	0,9999988	0,9999988	0,9999989	0,9999989	0,9999990	0,9999990	0,9999991	0,9999991	0,9999992
4,8	0,9999992	0,9999992	0,9999993	0,9999993	0,9999993	0,9999994	0,9999994	0,9999994	0,9999995	0,9999995
4,9	0,9999995	0,9999995	0,9999996	0,9999996	0,9999996	0,9999996	0,9999996	0,9999997	0,9999997	0,9999997
5,0	0,9999997	0,9999997	0,9999997	0,9999998	0,9999998	0,9999998	0,9999998	0,9999998	0,9999998	0,9999998

Abraham de Moivre (1667-1754)

Pierre-Simon de Laplace (1749-1827)

Johann Carl Friederich Gauss (1777-1855)