Concours d'Agrégation
Rapports des Jurys sur l'épreuve de Modélisation

Extraits intéressant les candidats de l'option "Probabilités"

1999

Epreuve de modélisation

La répartition selon les options a été de 314 candidats présentant l'option calcul scientifique, et de 218 présentant l'option probabilités statistiques (parmi les admissibles présents aux épreuves). Le choix du texte a été le fait de 136 candidats, soit environ le quart d'entre eux. Les notes obtenues ont été en moyenne un peu inférieures à ce qu'elles ont été dans le cas du choix de la leçon. Cette observation rend compte du fait que la connaissance des titres des leçons depuis un an et leur nombre relativement petit avait permis une étude exhaustive de celles-ci.

L'organisation matérielle de l'épreuve s'est montrée pertinente, aucun incident de nature à gêner les candidats n'est venu la perturber. Les nombreux observateurs extérieurs ont pu être accueillis dans de bonnes conditions. L'habileté très variable des candidats a été prise en compte par une aide diversifiée et efficace des membres du jury chargés de la permanence dans les salles de préparation.

Plus de la moitié des candidats a fourni une illustration sur machine, montrant ainsi leur volonté et celle des préparations de jouer le jeu dès la première année. Les illustrations ont été de qualité très hétérogène, mais elles signifiaient dans tous les cas la prise en main d'un logiciel au moins. L'usage de l'ordinateur a montré qu'il était possible de réaliser des présentations à la fois pertinentes et relativement simples. La virtuosité n'était en rien indispensable.

D'une façon générale, le jury a apprécié dans beaucoup de cas une préparation sérieuse et pertinente à cette nouvelle épreuve. Il y a eu une volonté réelle de jouer le j eu dans la plupart des préparations. Il reste néanmoins deux points sur lesquels des efforts doivent être portés pour les prochaines sessions :

* d'une part, les thèmes du programme n'ont pas été assez exploités, hormis le thème dynamique des populations qui a été un réel succès. Certains candidats ont déclaré (avec mauvaise foi, je présume), qu'ils n'en avaient pas entendu parler, et étaient incapables seulement de les citer ! La plupart des leçons et quasiment tous les textes n'ont pas donné lieu à des illustrations dans ces thèmes, au mépris du programme de l'épreuve.

* d'autre part, la notion de modélisation a semblé très peu claire pour bon nombre de candidats. Il serait bon qu'ils soient préparés à pouvoir expliquer pourquoi tel ou tel modèle est naturel, sous quel type de conditions il est valable etc ...

Ajoutons que certains candidats ont cru pouvoir recopier des programmes pillés dans les ouvrages spécialisés sur Maple, Matlab etc... Il faut signaler que le jury ne se prive pas d'interroger dans le détail sur la façon de procéder (il était permis d'utiliser un programme tout fait à condition de le comprendre) et que, en raison de cette déviance, certains de ces ouvrages ne seront plus autorisés. Si le travail d'illustration sur machine a été justement récompensé, le jury a mis de fort mauvaises notes aux virtuoses de l'ordinateur qui ne comprenaient rien aux mathématiques.

Remarques particulières sur les contenus

Le jury attire l'attention des candidats sur la nécessité, commune aux trois épreuves, de soigner l'exposition, la présentation et la qualité des explications.

La compréhension des ressorts de la statistique est souvent déficiente. Le programme n'est pas ambitieux, mais il importe que les candidats comprennent ce qu'ils font, par exemple quand ils construisent un intervalle de confiance ou un test, et qu'ils sachent distinguer, dans un contexte de statistique, paramètre supposé connu dans le modèle, paramètre inconnu et donnée aléatoire issue de l'observation.

Il serait bon pour la plupart des candidats d'éviter des développements trop techniques par rapport aux illustrations qu'ils peuvent en donner.

Malgré les incitations du programme, les applications autres qu'internes ne sont pas nombreuses.

Les exposés sur le calcul d'intégrales pourraient utiliser des suites équiréparties en parallèle avec les méthodes de Monte-Carlo. Il faudrait souligner le fait que ces dernières ne sont justifiées que si la fonction à intégrer est très irrégulière, et dans le cas de fonctions de plusieurs variables qui seraient très difficiles à intégrer numériquement sans cela. Dans le cadre de la méthode de Monte-Carlo, on peut mentionner l'intérêt de l'algorithme d'acceptation-rejet pour simuler des variables aléatoires lorsque les méthodes usuelles de simulation s'avèrent impraticables.

Les sujets de statistique ont été l'objet d'un rejet de la part de nombreux candidats qui ont alors préféré le texte.

Concernant le sujet "Loi des grands nombres... ", les candidats doivent penser à la consistance d'estimateurs numériques en statistique. Il faudrait aborder, même succinctement la question de l'amplitude des fluctuations (théorème central limite, inégalités exponentielles, loi du logarithme itéré) qui devrait être illustrée. Pour les candidats les plus avancés, il ne faut pas oublier les relations avec la théorie des martingales : cette théorie contient d'importants résultats de convergence presque sûre.

Concernant le sujet "Test du χ^2 pour lois multinomiales ", il faut penser et savoir utiliser si possible le test d'indépendance du χ^2. L'exemple des lois de Mendel est intéressant : pouvoir expliquer la raison des probabilités 9/16, 3/16, 3/16, 1/16 apparaît souhaitable. Il faudrait aussi savoir simuler des réalisations d'une loi multinomiale à partir de la fonction random. Il faudrait aussi dans la mesure du possible savoir comment utiliser les principes du test du χ^2 pour tester l'adéquation d'une famille paramétrée de lois à un ensemble d'observations.

Concernant le sujet "Construction d'intervalles de confiance", il faut comprendre que dans un cadre général, il existe à priori une infinité d'intervalles de confiance de niveau donné $a \in (0, 1)$, mais que l'on choisit celui qui est optimal en tel ou tel sens :

- A un niveau élémentaire, on peut imposer l'intervalle de longueur minimale

- En fait, le fait d'être optimal correspond à la notion de test uniformément le plus puissant.

Dans tous les cas les intervalles ne sont pas nécessairement symétriques. On peut utiliser l'exemple des sondages d'opinion.

Concernant le sujet “Théorème central limite”, il faut penser à la relation avec la construction d’intervalles de confiance, avec la fonction de répartition empirique, ainsi qu’avec le test du χ^2. On peut penser à l’approximation des lois de Poisson et Gamma. Il est intéressant aussi de considérer des variables aléatoires indépendantes, mais n’ayant plus la même loi. D’autres lois limites peuvent apparaître : par exemple la loi de Poisson dans le cas binomial ou les lois limites des extrêmes.

A propos de la *fonction de répartition empirique*, on aimerait voir précisé, au-delà du théorème de Glivenko-Cantelli, l’ordre de la vitesse de convergence vers zéro, quand n tend vers l’infini, de $\sup_x |F_n(x) - F(x)|$, qui est $O(1/\sqrt{n})$. Il faut aussi introduire, si possible, la réciproque généralisée d’une fonction de répartition. On peut alors présenter quelques résultats sur l’estimation de quantiles. Du point de vue de la modélisation, il est essentiel de s’intéresser aux tests de Kolmogorov- Smirnov ou Cramer- von Mises et, si possible, de les comparer à la version du test d’adéquation du χ^2, qui s’applique à des échantillons qui ne sont pas nécessairement multinomiaux. Il serait bon de faire la connexion avec le théorème central limite.

A propos du sujet “Chaînes de Markov à espaces d’états fini” : il faut aborder la question des points absorbants : y parvient-on en un temps fini presque sûrement ? Il faut connaître le rle du spectre de la matrice de transition. Il faut connaître le théorème de Perron-Frobenius. Il faut savoir que si l’espace d’états n’est plus fini, une chaîne de Markov irréductible peut être soit transiente, soit récurrente positive, soit récurrente nulle. Il faut savoir simuler une suite de variables aléatoires formant une chaîne de Markov à un nombre fini d’états à partir de “random”. Il faut connaître la loi des grands nombres relative à de telles chaînes de Markov homogènes à nombre fini d’états, irréductibles.

Les textes ont donné plus de mal aux candidats, même si un certain nombre d’entre eux ont présenté un exposé de qualité. Il est clair qu’ils y étaient moins bien préparés. Beaucoup de candidats n’avaient aucune idée de ce que l’on attendait d’eux et, par exemple, négligeaient de commencer par une présentation et une synthèse du texte, répondant directement aux questions alors que celles-ci n’avaient pour objectif que de donner des pistes à ceux qui se seraient sentis un peu démunis. Rappelons qu’il ne s’agit pas de faire une lecture servile, pas plus que de traiter un problème déguisé. Les candidats doivent avoir comme l’un de leurs soucis essentiels le fait de rassembler leurs connaissances autour du thème du texte choisi.

Conclusion

Outre l’intérêt évident de pousser les candidats à s’intéresser à la modélisation et au calcul effectif, le fait de pouvoir les interroger plus longuement a paru très positif, Le mode interactif les a mis vraiment en confiance. Le caractère moins formaliste de la nouvelle épreuve instituant un dialogue avec les candidats a permis au jury de juger beaucoup plus facilement du niveau mathématique de ceux-ci, ainsi que le déclarent en particulier par comparaison les membres du jury ayant participé à l’oral l’an dernier.

Les qualités comme la capacité à synthétiser, à reconnaître les phénomènes importants, à ouvrir son enseignement pour le rendre vivant et actuel sont plus naturellement décelables dans cette épreuve, et c’était bien là l’un de ses objectifs.

La maîtrise d’un logiciel de calcul scientifique sera confirmée l’an prochain, mais était déjà significative cette année, et c’est un point important dans le recrutement d’enseignants.

Ainsi, les trois épreuves se complètent et participent à la tentative de cerner la personnalité scientifique du candidat, ainsi que d'évaluer ses connaissances et son savoir-faire en mathématiques, afin de décider avec le plus grand discernement possible s'il est raisonnable de le recruter.

2000

Oral de Modélisation - Généralités

- *Modèles mathématiques*

L'objectif de l'épreuve de modélisation est l'étude de modèles mathématiques issus de domaines applicatifs variés. Dans une démarche d'enseignant, ceci impose de faire le lien entre les théories mathématiques et leur utilisation dans un cadre scientifique ou technique. Le jury apprécie donc tout particulièrement que le candidat s'efforce de mettre cette articulation en lumière, en construisant son exposé autour d'un véritable problème de modélisation.

Les propriétés que doit dégager la modélisation mathématique sont qualitatives et quantitatives. Figurent donc au programme des méthodes numériques, statistiques et symboliques. Toutefois, et contrairement aux deux autres épreuves d'oral, l'accent est mis ici sur l'application au modèle: choix et adéquation d'une méthode, limitations théoriques ou pratiques,... Le jury apprécie donc aussi bien que le candidat mobilise ses connaissances de base, que l'obtention de résultats quantitatifs par des méthodes numériques et statistiques.

- *Utilisation de l'outil informatique*

L'objectif de l'épreuve n'est pas la démonstration de la dextérité informatique, mais la maîtrise des logiciels au programme pour illustrer et soutenir une démarche scientifique et pédagogique. Ces derniers sont puissants et donnent facilement accès à de nombreuses méthodes.

Le jury a particulièrement apprécié les candidats capables d'expliquer leur mise en oeuvre informatique, voire de la modifier légèrement pour répondre aux questions. A contrario, l'incapacité à effectuer la moindre modification, le fait de ne pas être capable de mettre au point le programme de démonstration font mauvaise impression.

D'une manière générale, l'utilisation de l'outil informatique et des logiciels du programme dans le cadre d'une épreuve orale requiert:

1. une préparation avant le concours: il s'agit essentiellement de bien connaître des solutions simples pour obtenir et illustrer des résultats avec les logiciels au programme: tracé de courbes et surfaces, simulations de processus, solution numérique ou symbolique d'équations différentielles ordinaires ou aux dérivées partielles. Il faut connaître ses limites et savoir ce que l'on pourra produire à coup sûr en moins d'une heure de préparation.

2. une méthode de travail pendant les 4 heures de préparation précédant l'épreuve. Elle réservera le temps de produire des illustrations informatiques pertinentes. L'insertion de cette illustration dans l'exposé de modélisation pourra souvent se faire simplement en mettant en lumière l'effet de paramètres essentiels du modèle ou des méthodes de calcul.

- *Comportement face à un texte*

Le candidat doit s'efforcer de s'"approprier" le texte, de décider comment ce texte sera exploité dans l'exposé. Ceci passe par la constitution de notes personnelles. On constate que les candidats qui s'appuient uniquement sur le texte original lors de leur exposé sont contraints à une paraphrase. Ils ont du mal à l'exploiter en "temps réel" pour dégager les idées qu'ils veulent exprimer.

Trop de candidats traitent le texte "comme un problème", en dépit des mises en garde du préambule. Ceci n'est pas conforme à l'esprit de l'épreuve qui recherche l'exploitation du texte dans un exposé de modélisation mathématique. De la même façon, les développements suggérés doivent être considérés comme des pistes pour s'approprier le texte et l'exploiter. Leur usage est facultatif.

Le candidat a intérêt à relier les développements du texte aux outils mathématiques qu'il connait, à faire valoir sa culture scientifique et à faire preuve de curiosité intellectuelle. Ceci peut se faire par des rappels de définition, ou en précisant le cadre d'application des principaux théorèmes invoqués par le texte. Il faut cependant éviter les longues digressions hors sujet qui ne font progresser ni la compréhension ni l'exploitation du texte.

Oral de Modélisation - Option Probabilités

Remarque sur la statistique mathématique: cette discipline scientifique se caractérise par des problématiques propres associées à une modélisation fondée sur les probabilités. Il est nécessaire de bien établir la problématique et de lui associer la modélisation appropriée à l'occasion des leçons qui relèvent de cette discipline, que ce soit totalement (tests, test du χ^2, intervalle de confiance) ou partiellement (loi des grands nombres appliquée à la statistique, fonction de répartition empirique, propriétés asymptotiques et leur applications à la statistique, ...). La leçon sur les tests du χ^2 requiert, pour la clarté de l'exposition, de bien distinguer entre la loi du χ^2, la statistique du χ^2 associée à un échantillon discret, le quantile d'ordre a associé à la loi du χ^2, éventuellement la statistique du χ^2 associée, après discrétisation, à un échantillon de la loi non discrète. La confusion entre ces notions rend souvent incompréhensibles les exposés sur ce thème.

Illustration informatique dans l'option de probabilités et statistisques: On attend des simulations de Monte-Carlo, c'est à dire la réplication d'un nombre suffisant d'expériences à partir de variables aléatoires indépendantes; on met ainsi en évidence les fluctuations du phénomène modélisé. Les résultats peuvent conduire à des représentations sous forme d'histogrammes, de fonctions de répartition empiriques, des caractérisations numériques comme moyennes et écarts-types,... En définitive, il importe de mettre en évidence la loi de probabilité d'une variable ou d'un estimateur.

Les applications doivent être substantielles:

* Les jurys aimeraient voir moins de candidats calculer p par une méthode de Monte-Carlo. Le calcul d'intégrales par la méthode de Monte-Carlo doit être motivé et illustré par des exemples à plusieurs variables, dans une situation où les méthodes numériques sont inadaptées.

* Tester l'hypothèse qu'un dé à six faces n'est pas pipé, n'est pas un exemple captivant pour l'auditoire...

* Il en va de même du simple test que le générateur aléatoire de Maple fournit des pseudo–aléatoires uniformes sur [0, 1],

Les candidats ne doivent pas se borner à énoncer des propriétés qu'ils ne comprennent que superficiellement:

Lorsqu'on signale qu'un estimateur est sans biais, il faut également être capable d'expliquer qu'un estimateur $\hat{\theta_n}$ d'un paramètre θ est dit sans biais lorsque $E[\hat{\theta_n}] = \theta$.

Très peu de candidats ont compris ce que représentait la fonction de taux de défaillance en fiabilité. D'une manière générale, les notions associées au terme “taux” (de naissance, de mortalité, d'intérêt, de pannes) semblent mal comprises. Peu de candidats savent expliquer comment évolue en fonction de T l'espérance conditionnelle de vie d'un être vivant sachant qu'il a déjà vécu T unités de temps. La loi de la durée de vie d'un être vivant, considéré comme le représentant d'une population homogène, ne peut être modélisée par une loi exponentielle.

L'énoncé suivant, entendu cette année, laisse perplexe: ” La formule de Stirling est équivalente au théorème de la limite centrale.”

Des candidats trop nombreux confondent une variable aléatoire avec sa loi de probabilité. Quand on additionne des variables aléatoires indépendantes, on effectue le produit de convolution des lois correspondantes.

La méthode de simulation de Metropolis-Hastings, qui produit en régime stationnaire des variables aléatoires dont chacune suit une loi donnée, n'est pas appropriée pour simuler des lois que l'on sait simuler directement. Par exemple, elle n'est pas appropriée au cas où on veut produire une suite finie de variables aléatoires indépendantes de même loi multinomiale.

2001

Modélisation.

Le jury constate que les travers signalés dans le rapport de l'an dernier sont moins fréquents cette année. Les remarques faites dans le rapport de la session 2000 ne seront pas répétées ici, bien qu'elles restent d'actualité. Par ailleurs, le niveau des présentations informatiques s'est bien amélioré en moyenne.

Cloisonnement:

On constate durant l'oral de modélisation des manques surprenants, sans doute dus à une vision trop cloisonnée du domaine. Nous insistons sur ce point, car il requiert aussi bien un travail thématique en algèbre et analyse, qu'un entrainement à mobiliser ses connaissances en dehors de cadres bien délimités.

Modélisation:

Le jury apprécie que les candidats soient capables d'interpréter les modèles introduits et les résultats obtenus sur ces modèles en termes autres que mathématiques. C'est notamment vrai pour les textes, lesquels se rattachent parfois à des situations que l'on peut rencontrer dans la vie courante et pour lesquels la confrontation du modèle avec le phénomène réel n'est pas hors de portée.

Esprit critique:

L'esprit critique est une bonne chose et doit être encouragé chez les candidats, il leur est par contre inutile d'adopter une attitude systématiquement critique vis-à-vis des textes. Celle-ci ne se justifie en général pas, même s'il est vrai que certains points de certains textes peuvent donner lieu à discussion ou requièrent une interprétation. Les candidats devraient profiter des ouvertures dans les textes, et les exploiter. Lorsqu'ils emploient la phrase ”ils disent dans le texte”, en général en pensant avoir trouvé une faille ou un point obscur, qu'ils pensent que l'auteur est peut être sur le banc du jury.

Méthodes Numériques:

Ce n'est pas parce que la formule de Black et Scholes provient d'une modélisation stochastique que la meilleure méthode pour évaluer l'intégrale est la méthode de Monte-Carlo. La comparaison entre méthode numérique classique et méthode de Monte-Carlo est appréciée à chaque fois que le sujet d'y prête.

Vecteur aléatoire:

A propos d'une leçon sur la "simulation d'une variable ou d'un vecteur aléatoire", le jury aimerait que la partie relative aux vecteurs soit plus souvent traitée. En particulier ceci contient le cas de vecteurs aléatoires gaussiens de moyenne et de matrice de variance-covariance donnée. Il peut être utile d'envisager le cas ou cette matrice n'est pas de rang maximal.

Modèles:

Les candidats devraient savoir expliquer la relation entre le contexte gaussien et le méthode des moindres carrés et le modèle linéaire.

Lois:

On aimerait que les candidats sachent expliquer l'emploi des lois exponentielle et de Poisson. On peut souhaiter qu'ils puissent par exemple réagir devant des questions comme: Pourquoi la durée de vie d'une lampe à incandescence habituelle peut être approximativement représentée par une loi exponentielle? Comment faire si le filament de la lampe s'évaporait peu à peu lorsque la lampe est allumée?

Monte-Carlo:

* Le calcul du volume de la boule unité de R^d n'a guère d'intérêt, la formule explicite étant connue. Le volume d'un convexe décrit par la solution d'un problème de "Programmation Linéaire" en dimension supérieure à 4 serait plus intéressant à calculer. Les méthodes de Monte-Carlo permettent également d'approcher un tel convexe par essais et erreurs.

* En ce qui concerne le calcul des intégrales, il serait bon qu'on s'intéresse également à des intégrales du type:

$$\int_{[0,1]^d} f(x_1, \ldots, x_d) g(x_1, \ldots, x_d) dx_1 \cdots dx_d,$$

dans lesquelles g est une densité de probabilité sur R^d.

* Dans toutes les applications des méthodes de Monte-Carlo, on souhaite la détermination d'un intervalle de confiance, la visualisation des fluctuations des résultats obtenus par des tirages indépendants. L'intervalle de confiance est l'analogue de la majoration d'erreur recherchée dans les méthodes d'intégration numérique non probabiliste. Lorsque le contexte s'y prête, il est souhaitable que les candidats sachent comparer l'efficacité des méthodes numériques non probabilistes avec celles de Monte-Carlo.

Théorème de la Limite Centrale:

Il serait bon que les candidats aient des idées sur ce qui se produit lorsque les variables aléatoires n'admettent pas de moment d'ordre 2. L'exemple des lois de Cauchy est particulièrement instructif. Cela peut faire l'objet de simulations et de représentations. La question analogue se pose pour les chaînes de Markov.

Tests d'hypothèses:

Lorsqu'on ne connaît pas d'expression utilisable pour la loi de la statistique de tests sous H0, on peut la simuler. Ceci est également vrai lorsqu'on dispose de la loi limite sous H0 quand la taille de l'échantillon tend vers l'infini, mais que l'échantillon dont on dispose est de petite taille.

Générateur pseudo-aléatoire:

La question du test d'un générateur pseudoaléatoire est intéressante. Il faudrait savoir différencier une suite de nombres pseudo-aléatoires, par exemple équirépartis sur [0, 1] , d'une suite de variables aléatoires uniformes sur [0, 1].

2002

Epreuve orale de Modélisation.

D'une manière générale, on constate que les candidats sont maintenant bien préparés à cette épreuve. Le nombre de candidats faisant le choix d'un texte a crû.

Toutefois, l'épreuve reste exigeante car elle demande un certain recul par rapport aux connaissances acquises et met particulièrement en avant l'utilisation de résultats en "aval" des développements théoriques usuels.

On doit conseiller aux candidats de mieux faire la distinction entre le processus de "Modélisation" et celui de "Simulation". Toutefois, la discussion de propriétés quantitatives des modèles aboutit très fréquemment à une simulation qu'il s'agit alors de bien conduire.

Probabilités et Statistiques.

Les textes et leçons faisant intervenir des vecteurs gaussiens (lois normales multidimensionnelles) se sont révélés un test exigeant des connaissances en algèbre linéaire des candidats, et de leur capacité à les appliquer. Les points suivants en sont des exemples:

* réduction de formes quadratiques en relation avec la simulation de vecteurs gaussiens dont la matrice de covariance est donnée,

* utilisation de la réduction des formes quadratiques, et plus précisément de la détermination des axes d'un ellipsode, dans l'analyse en composantes principales d'une loi normale multidimensionnelle,

* utilisation des notions liées à la projection orthogonale, par exemple dans les méthodes de moindres carrés et celles d'analyse en composantes principales.

* utilisation de la réduction des matrices et plus particulièrement du théorème de Perron-Frobenius dans l'étude des chaînes de Markov à espace d'états fini.

Il serait également utile que les candidats connaissent l'intérêt de méthodes probabilistes dans l'étude de questions purement mathématiques. Par exemple:

* accélération de la recherche de grands facteurs premiers, par exemple en relation avec la cryptographie,

* intérêt historique et heuristique du modèle probabiliste de Cramer de la suite des nombres premiers: on tire successivement au hasard des variables aléatoires de Bernoulli de paramètres $p_n \simeq 1/\log(n)$ quand $n \to \infty$, les pseudo nombres premiers sont les instants auxquels on obtient 1.

2003

Epreuve orale de Modélisation.

D'une manière générale, on constate que les candidats sont maintenant bien préparés à cette épreuve. Toutefois, l'épreuve reste exigeante car elle demande un certain recul par rapport aux connaissances acquises et met particulièrement en avant l'utilisation de résultats en aval des développements théoriques usuels. On doit conseiller aux candidats de mieux faire la distinction entre le processus de Modélisation et celui de Simulation . Toutefois, la discussion de propriétés quantitatives des modèles aboutit très fréquemment à une simulation qu'il s'agit alors de bien conduire.

Probabilités et Statistiques.

Il est indispensable que tous les candidats qui choisissent cette option aient assez de recul pour pouvoir expliquer de manière convaincante les notions de base. Par exemple ce qu'est une variable aléatoire. Concernant les chaînes de Markov, il importe que les candidats sachent faire la classification des états, sur des exemples simples. Le jury a constaté cette année encore des réponses absurdes. Les préparateurs aideront les candidats en leur présentant des exemples concrets simples. Parmi les notions mal maîtrisées, figurent également les tests statistiques, l'utilisation dans le contexte probabiliste des convolutions et fonctions génératrices. Les textes et leçons faisant intervenir des vecteurs gaussiens (lois normales multidimensionnelles) se sont révélés un test exigeant des connaissances en algèbre linéaire des candidats, et de leur capacité à les appliquer. Les points suivants en sont des exemples:

* réduction de formes quadratiques en relation avec la simulation de vecteurs gaussiens dont la matrice de covariance est donnée,

* utilisation de la réduction des formes quadratiques, et plus précisément de la détermination des axes d'un ellipsode, dans l'analyse en composantes principales d'une loi normale multidimensionnelle,

* utilisation des notions liées à la projection orthogonale, par exemple dans les méthodes de moindres carrés et celles d'analyse en composantes principales.

* utilisation de la réduction des matrices et plus particulièrement du théorème de Perron–Frobenius dans l'étude des chaînes de Markov à espace d'états fini.