



INSEE
DIRECTION GÉNÉRALE
INSTITUT NATIONAL
DE LA STATISTIQUE
ET DES ÉTUDES
ÉCONOMIQUES

SECRETARIAT GENERAL
Département des ressources humaines
Division "Gestion statutaire"
Section concours et examens

Rapport du jury du concours d'Administrateur Interne 2011

18 candidats se sont **inscrits** au concours interne d'administrateur 2011.

12 candidats se sont présentés à l'ensemble des **épreuves écrites**.

7 candidats ont participé aux **épreuves orales**.

1. Épreuve de résumé de texte

1.1. Épreuve de résumé de texte (durée 3 heures - Coef.15)

L'épreuve consistait cette année en la présentation synthétique du premier chapitre de La société des individus de Norbert Elias en 500 mots, une marge de +/- 10% étant admise, le tout en 3 heures. Comme l'année précédente, où un extrait d'Amartya Sen avait été proposé aux candidats, il s'agissait d'un texte classique des sciences sociales, proche des possibles occupations professionnelles de candidats au concours d'administrateur de l'INSEE.

Cette épreuve demandait aux candidats de rendre l'essentiel du texte, (concepts clés et définitions, démarche argumentative), tout en conservant la dynamique de la pensée et en respectant le sens du texte sous la contrainte d'un nombre de mots fixés. Le jury a été surpris par le nombre d'erreurs de compréhension et la fréquence des faux sens, voire des contresens dans certaines copies. Norbert Elias s'interroge sur ce qu'est une société et la thèse de l'auteur, comme l'articulation de sa pensée, n'était parfois absolument pas comprise. Ces erreurs nous ont paru plus fréquentes que les années précédentes.

De même, le niveau de français n'était pas toujours à la hauteur de ce qu'on est en droit d'attendre de candidats à un tel concours. Il a été cette année encore assez moyen au vu de l'ensemble des copies. Les constructions des phrases sont trop souvent grammaticalement incorrectes et les fautes de conjugaison des verbes encore trop nombreuses. Le jury rappelle l'importance de la relecture, fondamentale dans l'écriture d'un résumé.

À ces fautes de français se sont ajoutées des erreurs de construction, des présentations trop monolithiques ou, au contraire, marquées par une utilisation du « passage à la ligne » trop fréquente pour avoir une signification opératoire et des écarts importants vis-à-vis du texte proposé au résumé. Il est également rappelé aux candidats que la paraphrase du texte est à bannir. Ainsi, il ne suffit pas de reprendre sans explicitation le vocabulaire utilisé (et défini) dans le texte, voire de coller des citations les unes aux autres. De même, un bon résumé ne reprend pas obligatoirement la structure du texte.

Si cette année tous les candidats ont pris soin de respecter la contrainte des 500 mots à +/- 10 %, une copie était inachevée, le candidat n'ayant pas eu le temps de recopier intégralement ses brouillons. C'est d'autant plus dommageable que la partie dont le jury a eu connaissance était particulièrement réussie. Les candidats sont invités à prêter une attention particulière à la gestion de leur temps, qui fait partie des contraintes de l'épreuve.

Finalement, les copies fidèles à l'extrait présenté, à l'orthographe irréprochable, à l'expression maîtrisée et répondant à la consigne ont été très peu nombreuses et ont été particulièrement valorisées par le jury. Comme les années précédentes, l'épreuve de résumé s'est révélée particulièrement discriminante entre les candidats, ce qui a permis d'utiliser assez largement l'éventail des notes.

1.2. Épreuve d'exposé (durée 30mn - Coef.10)

Les candidats disposaient de 3 heures de préparation sur un dossier portant sur un thème d'actualité socio-économique d'une cinquantaine de pages. Ils avaient quinze minutes pour proposer un exposé retraçant les enjeux du dossier fourni. Dix à quinze minutes étaient ensuite consacrées à une discussion partant des questions abordées dans le dossier, mais ne s'y limitant pas.

Les sujets proposés cette année étaient tous de difficulté égale et étaient aisément traitables par des candidats curieux de l'actualité et des faits économiques et sociaux. Dans



l'ensemble, les candidats ont respecté le temps qui leur était imparti et ont proposé une restitution convenable des éléments contenus dans le dossier, même si les prestations étaient en moyenne moins bonnes que les années précédentes. On regrettera en particulier des problèmes criants de méthode chez certains candidats. Il s'agit pour le candidat de tirer de l'ensemble de la documentation fournie les éléments les plus importants et d'en saisir les enjeux.

Le temps de préparation doit permettre au candidat de problématiser et d'organiser sa présentation de manière à ne pas se limiter à une simple accumulation des points traités dans les différents documents. Il faut insister sur le fait que les dossiers peuvent être longs à lire au regard du temps disponible (voire denses) et qu'il n'est donc pas question d'en faire un résumé exhaustif, un simple commentaire, ou une "synthèse". C'est pourquoi la prestation doit être explicitement organisée et non simplement discursive : exposé d'une problématique induite par la nature du dossier puis d'une démarche exprimée par un plan, qui peut être simple mais doit avoir été clairement formulé.

Dans l'exposé, on attend une vraie introduction problématisée, un plan clair et progressif et une conclusion. En revanche, il n'est pas nécessaire en introduction de présenter un à un les documents qui peuvent être très nombreux. On rappellera également qu'il s'agit dans l'exposé de s'appuyer sur les documents fournis et qu'une présentation qui n'y ferait pas appel et présenterait une vague thèse personnelle sans lien avec les travaux et articles proposés est à proscrire.

L'entretien consécutif à l'exposé départageait assez nettement les admissibles. Le jury insiste sur le rôle très important des questions dans une telle épreuve : il faut que les candidats accordent un grand soin à ces questions et ne croient pas que tout est joué après l'exposé. L'absence de culture historique ou des repères chronologiques fantaisistes notamment sont du plus mauvais effet, même si l'exposé a été correctement mené auparavant. Le jury a par ailleurs été surpris que certains candidats à un poste d'administrateur de l'INSEE se laissent déstabiliser par des questions, simples, sur la construction des chiffres cités dans le dossier. Cela fait également partie de la culture générale attendue chez les candidats à un tel concours.

2. Épreuve de statistiques

2.0. Introduction

2.0.1. Politique de mise en compétition

2.0.1.1. Cette année a vu l'introduction, dans le texte des sujets, de deux notes de mise au point. La première, intitulée « organisation de l'épreuve » avait pour but de présenter un minimum d'informations générales sur les sujets proposés, de façon à « prévenir » les candidats du contenu des sujets et, par là-même, de définir leur stratégie en fonction de leurs « affinités » relativement aux sujets.

La seconde constituait un ensemble de préalables destiné à fixer les notations utilisées dans les énoncés, ce afin de prévenir des ambiguïtés éventuelles. Des rappels succincts de définitions (cours) étaient indiqués, dans un but de « normaliser » les pré-supposés entre candidats.

2.0.1.2. La correction de l'épreuve de statistique a suivi un principe de stricte conformité au programme : les concepts ou méthodes explicitement hors programme, mais utilisés par des candidats pour résoudre les questions posées, ont été considérés comme non venus.

Cependant, une telle attitude ne laisse guère de latitudes aux examinateurs en matière de choix de sujets (hormis les simples questions de cours ou leurs variantes).

Afin de ne pas pénaliser certains candidats mieux formés en statistique, on a donc posé certains des sujets en utilisant une démarche progressive. Ainsi, la méthode d'estimation par maximisation de la vraisemblance ne figure pas au programme : cependant, les connaissances nécessaires pour accéder et appliquer cette méthode (échantillonnage, loi de



probabilité produit dans le cas iid, maximisation de fonction) en font clairement partie. On a donc, non seulement dans les préalables, mais aussi dans les sujets concernés eux-mêmes, défini la fonction de vraisemblance et jaugé l'aptitude des candidats à exhiber l'estimateur attendu.

2.1. Écrit

Deux problèmes et deux exercices ont été proposés, équilibrés afin de combiner des parties « cours et technicité » et des parties « réflexion et initiative ». L'ensemble a donc balayé une très large partie du programme, afin d'éviter au maximum des effets d'aubaine (positifs ou négatifs) entre candidats.

2.1.1. Le contenu des sujets était le suivant.

L'exercice 1 était sans grande surprise : il portait essentiellement sur des calculs de probabilités d'événements, sur une normalisation de densité, sur un schéma d'urne et sur la démonstration du théorème de Poincaré.

L'exercice 2 portait sur l'étude d'une loi de probabilité discrète (dite « tri-hypergéométrique »), avec, en préliminaire, la démonstration d'une identité classique (celle de Van der Monde) qui ne supposait que la connaissance de la formule du binôme.

Le problème 1 comparait trois estimateurs d'un même écart-type (moyenne absolue, racine carrée de moyenne d'ordre 2 et maximum de vraisemblance), et en étudiait les propriétés (biais, modes de convergence probabilistes).

Enfin, le problème 2, un peu plus « abstrait », portait sur l'estimation par intervalle de confiance, à partir de l'exemple de la loi de Bernoulli. Dans cet exemple élémentaire, la définition de la fonction génératrice a été posée et le candidat devait en étudier les propriétés (démonstration d'inégalités-types et de majorations, calcul de loi, convergence en probabilité). L'application portait sur le contrôle de qualité. Le problème se terminait sur une estimation par intervalle du paramètre de la loi de Bernoulli.

2.1.2. Dans l'ensemble, les résultats de l'écrit ont été très moyens, notamment par comparaison avec ceux de 2009 et 2010.

Seules deux excellentes copies se détachent, mais aussi deux copies quasi-insignifiantes. Le groupe « médian » est très fourni (environ la moitié des candidats). La distribution est donc à peu près tri-modale.

Parmi les défauts des candidats, on peut relever les suivants :

* deux derniers candidats : réponses insignifiantes, ou incohérentes, erreurs manifestes, bonne amorce parfois

* deux meilleurs candidats : réponse parfois peu explicite (mais juste), erreur de calcul avec résultat exact (ou erreur sans conséquence), oubli d'une identité de définition (eg $p + q + r = 1$), confusion entre loi du triplet (X, Y, Z) et celle du couple (X, Y) (marginalisation d'une loi), confusion entre convergence et non-biais, difficulté à amorcer un calcul de récurrence, intervalle de confiance « centré » sur l'estimateur (!) et non sur le paramètre à estimer ;

* autres candidats : démonstrations fausses (ou laborieuses, ou peu claires), équation de l'EQM fautive, bon résultat après un calcul d'espérance faux ou un raisonnement faux, confusion entre indépendance et disjonction, ou entre marginal et conditionnel, égalité de Koëning-Huygens fautive, ignorance de la notion de « sans biais » ou d'intervalle de confiance, système linéaire exact mais résolution fautive, erreur sur définition de la loi du chi-deux, confusion entre continuité (classe C^1) et intégrabilité, Student pour un chi-deux. Un candidat semblait avoir des capacités (mais son résultat global reste médiocre) ; un autre ne sait pas calculer une espérance.



Cependant, dans l'ensemble des candidats, on peut relever de bonnes intuitions ou des débuts de résolution valables mais sans suite, ou encore suivis de calculs avortés.

Il est bon de rappeler qu'une présentation et une rédaction ne peuvent qu'aider au raisonnement et faciliter l'appréciation du niveau d'aisance du candidat. Utiliser le théorème de la récurrence ou mener à bien un calcul d'intégrale sont des nécessités.

2.2. Oral

2.2.1. Le **contenu des sujets** est synthétisable comme suit.

1	comparaison de 2 populations	test de proportions, test de moyennes
2	convexité & biais d'un estimateur	régression avec régresseur stochastique
3	régression	estimation, test
4	loi de Weibull & de Chi-deux	fonction d'échec, estimateur du mv & IC
5	loi Log-normale	intervalle de prévision, mv & IC
6	loi de Pareto & de Chi-deux	estimateur du mv, biais, cv, IC sym.
7	loi géométrique	estimation et propr., cv, efficacité relat.
8	sondage ds une population à 2 catégories	lois & estimateurs, propriétés & compar.
9	jeu de cartes truqué (2 as de pique)	proba. de détection, loi du rang d'arrivée

Note : mv = maximum de vraisemblance, cv = convergence, IC = intervalle de confiance

Le déroulement de l'oral consiste, en bref, en l'exposé de la résolution par le candidat, suivi de quelques questions en vue de le « jauger au-delà » du sujet traité.

Dans l'ensemble, les candidats manquent de « hauteur de vue », ou de maturation, par rapport à la statistique. Est-ce l'enseignement de la statistique qui est trop « calculatoire » et trop peu « conceptuel », trop « procédural » et trop peu « méthodologique » ? Ainsi, trop de candidats se lancent trop rapidement dans les calculs, sans recul apparent vis-à-vis de la question posée (stress ?).

Ainsi, lorsqu'on leur demande d'indiquer les *grandes méthodes d'estimation existantes*, un certain nombre connaît (a) la méthode du maximum de vraisemblance, mais rares ont été ceux connaissant (b) la méthode des moments (du moins sous cette dénomination), qui est pourtant une méthode de base (et historiquement la première utilisée). Partant, la question sur leur connaissance (c) d'une troisième catégorie de méthodes, celles à distance minimale (moindres carrés ou moindres normes dans L_p , estimation non paramétrique de densités) n'a même pas été posée.

Un certain flou apparaît parfois dans la connaissance des concepts mis en jeu : (a) modes de convergences stochastiques, (b) écart quadratique moyen et arbitrage entre biais et dispersion, (c) association du couple « moyenne, écart-type » avec la loi normale, sans vision d'un équivalent dans le cas (eg) d'une loi dissymétrique ou multimodale.

De même, l'interprétation des paramètres d'une loi de probabilité n'est pas toujours claire pour certains candidats. Ainsi, l'un des 3 paramètres de la loi de Weibull peut (à une exponentiation près) s'interpréter comme un paramètre d'échelle, donc comparable à un écart-type, etc. Or, pour les 2 candidats ayant à traiter du sujet, cette compréhension n'était pas apparue d'emblée.

Des distinctions a priori élémentaires, comme (a) « théorique - empirique », (b) « paramètre - observation », (c) « distance finie - asymptotiquement », etc, semblent peu évidentes dans les esprits. Ainsi, le parallèle entre la définition d'une espérance (eg calcul d'intégrale avec la fonction de répartition théorique) et celle d'une moyenne arithmétique (simple) (calcul d'intégrale avec la fonction de répartition empirique) est parfois apparue comme une distinction lointaine.

Certains candidats récitent par cœur des réponses sans analyser réellement le contenu de la



question ou expliquent à l'aide de notions qu'ils ne maîtrisent pas. Il ne s'agit pas d'un oral de culture générale mais bien d'une photo des connaissances réellement connues et que le candidat est en mesure d'appliquer. Le manque de modélisation est parfois dommageable : illustrer une loi de Poisson, une loi géométrique ou une loi connue est un bagage minimum.

3. Épreuve écrite de mathématiques

La partie Analyse portait cette année sur l'étude des solutions d'équations différentielles linéaires d'ordre 1 ou 2. L'objet du problème était d'établir, à partir des définitions de base, plusieurs résultats classiques et importants concernant l'existence, l'unicité, la structure algébrique ou la périodicité des solutions. Il était précisé qu'aucune connaissance préalable sur les équations différentielles n'était nécessaire, les méthodes mises en œuvre étant celles de l'analyse classique présentes dans le programme du concours : dérivation / intégration, majoration d'intégrales, convergences de suites et séries...

Les deux premières parties étaient assez faciles et très détaillées : le candidat devait suivre les calculs demandés. La seule subtilité consistait en l'utilisation du Wronskien pour résoudre une équation du 2nd ordre et en les conditions d'annulation de celui-ci. Ces parties ont été traitées dans un bon nombre de copies. Les deux parties 3 (théorème de CAUCHY-LIPSCHITZ) et 4 (périodicité des solutions, en relation avec l'étude des valeurs propres d'un certain endomorphisme) étaient un peu plus difficiles et ont permis de départager les meilleurs candidats.

La partie Algèbre commençait par une première partie classique sur les polynômes d'endomorphismes suivie d'une seconde dans laquelle était prouvé un résultat de diagonalisation simultanée pour une famille d'endomorphismes diagonalisables qui commutent.

Le problème se terminait par deux courtes applications.

Dans l'ensemble, le sujet a été très discriminant et nous avons obtenu deux très bonnes copies, une bonne copie, cinq copies moyennes et quatre copies plus faibles (dont deux strictement inférieures à 5). On notera une copie complètement blanche en algèbre et quelques copies ayant obtenu une note assez mauvaise en analyse, partiellement compensée par une note honorable en algèbre.

Il nous a semblé aussi que les copies étaient meilleures que celles rencontrées lors des deux dernières sessions du concours

Les examinateurs ont toutefois noté beaucoup de confusions dans la manipulation des polynômes d'endomorphismes. Certains candidats n'ayant pas hésité, pour prouver les résultats demandés par l'énoncé, à écrire des produits de vecteurs, ou à utiliser des périphrases. Évidemment, les correcteurs ne peuvent en aucun cas valider des réponses aussi vagues et souvent erronées. En analyse, des imprécisions ont été notées dans l'écriture de primitives (intégrales indéfinies + constante, alors que l'énoncé demandait des intégrales définies).



ANNEXE

DONNEES STATISTIQUES

Évolution du nombre de candidats de 2000 à 2011

Année	Inscrits	présents	taux de présence	admissibles	taux admissibilité	admis	taux admission
2000	16	13	81,3%	8	61,5%	4	30,8%
2001	21	19	90,5%	9	47,4%	4	21,1%
2002	18	17	94,4%	8	47,1%	4	23,5%
2003	19	17	89,5%	9	52,9%	4	23,5%
2004	20	14	70,0%	10	71,4%	4	28,6%
2005	17	17	100,0%	8	47,1%	4	23,5%
2006	18	17	94,4%	8	47,1%	4	23,5%
2007	17	16	100,0%	7	41,2%	4	25,0%
2008	20	20	100,0%	10	50,0%	4	20,0%
2009	23	20	95,7%	8	36,4%	4	20,0%
2010	19	15	78,9%	10	66,7%	4	26,7%
2011	18	12	66,7%	7	58,3%	4	33,3%

Données sur les candidats par sexe

Année cumulée période 2000-2011	sexe	Inscrits	présents	taux de présence	admissibles	taux admissibilité	admis	taux admission
	Hommes	173	160	92,5%	79	45,7%	37	23,1%
Femmes	53	43	81,1%	23	43,4%	11	25,6%	

Provenance des candidats admis

CANDIDATS	2008	2009	2010	2011
ADMIS	4	4	4	4
dont hommes	4	4	3	4
dont femmes	0	0	1	0
HORS INSEE	3*	2*	1**	2*
INSEE	1	2	3	2

* Éducation nationale

** DGI

