

Statistical Methods for the Social Sciences

by Alan Agresti · Language: FR

Summary by booksummaryai.com —

<https://booksummaryai.com/country/fr/statistical-methods-for-the-social-sciences-resume>

Summary

En 60 secondes

Agresti enseigne les outils statistiques que les chercheurs en sciences sociales utilisent réellement : décrire les données, puis tester si les régularités sont réelles ou relèvent du bruit. Vous apprenez l'estimation, les tests d'hypothèses, la régression et l'analyse de variance comme une boîte à outils cohérente plutôt que comme des formules isolées. Les données catégorielles — effectifs, proportions, tableaux de contingence — reçoivent une attention particulière, reflet de la spécialité d'Agresti lui-même. L'ossature du livre tient en une seule discipline : quantifier l'incertitude avant toute affirmation.

Pourquoi le lire

Si vous lisez des sondages, menez des enquêtes ou évaluez des programmes, ce livre vous montre comment distinguer le signal du hasard. Il convient aux étudiants en sciences sociales, aux analystes des politiques publiques et aux autodidactes des données qui veulent de la rigueur sans démonstrations de théorie de la mesure. Les exemples d'Agresti proviennent de vraies questions de recherche, de sorte que chaque méthode arrive attachée à un problème auquel vous pourriez réellement être confronté.

Les idées clés

Commencez par les données que vous pouvez voir

Un étudiant ouvre le livre en s'attendant à des formules et reçoit à la place une question : que signifient ces nombres ? Agresti commence par la description — histogrammes, moyennes, médianes, écarts-types — en s'appuyant sur de vraies données en sciences sociales portant sur des sujets comme le revenu et l'éducation. Avant toute inférence, vous apprenez à résumer un échantillon désordonné en quelques nombres honnêtes. Le bénéfice vient quand ces résumés cessent d'être des exercices d'arithmétique et deviennent la base factuelle de toute affirmation ultérieure.

La description n'est pas un numéro d'échauffement ; c'est le fondement de l'inférence. Une moyenne sans notion de dispersion ne vous apprend presque rien, et un graphique peut révéler une structure qu'aucun test ne signalera. L'ordre choisi par Agresti a son importance : vous ne pouvez pas juger si un échantillon est inhabituel tant que vous ne savez pas à quoi ressemble le typique. Sauter directement aux valeurs p , c'est comme corriger un examen sans lire les

réponses.

À essayer: Prenez n'importe quel jeu de données que vous avez et calculez la moyenne et la médiane côte à côte ; si elles diffèrent beaucoup, tracez un histogramme avant de croire l'une ou l'autre.

La probabilité comme langage de l'incertitude

Le livre passe de la description d'échantillons au raisonnement sur les populations, et la probabilité devient le pont. Agresti travaille les distributions d'échantillonnage — ce qui se passe quand on tire des échantillons répétés de la même population — si bien que la courbe normale cesse d'être un ornement et commence à travailler. Le théorème central limite arrive comme la raison pour laquelle les moyennes se comportent de façon prévisible même quand les observations individuelles ne le font pas. Au terme de cette section, le hasard a une grammaire.

Les distributions d'échantillonnage sont l'idée la plus importante de la statistique introductive, et la plus difficile à saisir. Une fois que vous voyez une statistique d'échantillon comme un tirage de sa propre distribution, les intervalles de confiance et les tests de signification cessent d'être des rituels. La distribution normale importe moins parce que les données sont normales que parce que les moyennes de nombreuses observations y convergent. Cette convergence est ce qui permet aux sciences sociales de formuler des affirmations à partir d'échantillons modestes.

À essayer: Simulez 1 000 moyennes d'échantillon à partir de n'importe quel jeu de données, même avec un tableur, et regardez l'histogramme des moyennes se resserrer autour de la vérité.

Les intervalles de confiance : une plage, pas un verdict

Agresti construit l'intervalle de confiance comme le cheval de bataille de l'estimation : une estimation ponctuelle plus une marge d'erreur, avec 95 % de confiance comme norme des sciences sociales. Il montre l'intervalle pour une moyenne, puis pour une proportion, et insiste sur ce que le niveau de confiance promet réellement à travers des échantillons répétés. L'enjeu est pratique — un rapport de sondage qui ne donne qu'un pourcentage, sans intervalle, vous dit moins qu'il n'y paraît.

Un intervalle de confiance est une affirmation sur la méthode, pas sur un intervalle : utilisez la procédure de nombreuses fois et elle capture le paramètre 95 % du temps. C'est plus subtil que la lecture populaire selon laquelle un intervalle donné a 95 % de chances de contenir la vérité. Rapporter des intervalles à côté des estimations force l'honnêteté sur la précision, et les intervalles larges sont un avertissement, pas un échec. L'estimation répond généralement mieux à la question scientifique qu'une simple décision accepter-ou-rejeter.

À essayer: La prochaine fois que vous voyez un chiffre de sondage, exigez la marge d'erreur et calculez l'intervalle vous-même avant de tirer une conclusion.

Les tests de signification et leurs tentations

Ici, Agresti parcourt la logique en cinq étapes d'un test de signification — hypothèses, hypothèses statistiques, statistique de test, valeur p, conclusion — en utilisant des exemples comme le test de savoir si une nouvelle méthode d'enseignement modifie les scores. L'hypothèse nulle est présentée comme la position du sceptique : pas d'effet tant que les preuves ne disent pas le contraire. Il montre comment la valeur p mesure la surprise sous cette hypothèse nulle, et comment une petite valeur p est un signal, pas une preuve.

Une valeur p est la probabilité de données au moins aussi extrêmes si l'hypothèse nulle était vraie ; ce n'est pas la probabilité que l'hypothèse nulle soit vraie. La signification statistique répond à la question de savoir si un effet est distinguable du bruit, jamais s'il importe. Agresti pousse les lecteurs à rapporter les tailles d'effet à côté des valeurs p, car un effet trivialement petit peut être hautement significatif avec un échantillon assez grand. Mal lire ceci est l'erreur statistique la plus courante en sciences sociales.

À essayer: Pour tout résultat significatif que vous rencontrez cette semaine, demandez-vous quelle est la taille d'effet estimée — et si quelqu'un s'en soucierait à cette taille.

Comparer deux groupes

Le livre aborde la comparaison qui sous-tend la plupart des recherches en sciences sociales : le groupe A est-il différent du groupe B ? Agresti couvre le test t pour échantillons indépendants appliqué aux moyennes et la comparaison de deux proportions, avec des intervalles de confiance pour la différence. Les exemples traversent des dispositifs traitement-témoin et des contrastes démographiques. Le résultat est une boîte à outils pour la question que pose toute étude : quelque chose a-t-il réellement changé, ou en a-t-il seulement l'apparence ?

Comparer des groupes exige de tenir compte de deux sources d'aléa à la fois, si bien que l'erreur type d'une différence combine la variabilité des deux échantillons. Les échantillons indépendants et les paires appariées appellent des analyses différentes, et les confondre fausse votre incertitude. Une différence de proportions, accompagnée d'un intervalle, est souvent plus informative que l'une ou l'autre proportion prise isolément. La plupart des débats de politique publique portent secrètement sur des différences, non sur des niveaux.

À essayer: Quand vous lisez une affirmation selon laquelle un groupe surpasse un autre, vérifiez si la différence rapportée s'accompagne d'une erreur type ou d'un intervalle.

Données catégorielles et test du khi-deux

Voici le terrain de prédilection d'Agresti : les tableaux de contingence, où lignes et colonnes sont des catégories et où les cellules contiennent des effectifs. Il construit le test du khi-deux d'indépendance pas à pas — les effectifs attendus sous l'indépendance, la statistique qui additionne les écarts au carré, les degrés de liberté déduits de la forme du tableau. Un exemple comme l'appartenance partisane selon le sexe montre comment le test signale une association sans dire en quoi elle consiste.

Le khi-deux teste si deux variables catégorielles sont indépendantes, mais un résultat significatif n'est qu'un commencement : il faut inspecter les résidus standardisés pour repérer quelles cellules portent l'association. Le test exige des effectifs attendus suffisants dans les cellules, une condition que les étudiants négligent de façon routinière. Une association dans un tableau n'est pas une causalité, et des variables latentes peuvent fabriquer des motifs. Ce chapitre est la graine des travaux ultérieurs, déterminants pour sa carrière, qu'Agresti consacrera à l'analyse des données catégorielles.

À essayer: Croisez deux variables catégorielles qui vous tiennent à cœur, calculez les effectifs attendus, et vérifiez qu'aucune cellule ne passe sous le seuil habituel avant de vous fier au khi-deux.

La régression : un prédictor, mille usages

La régression linéaire simple entre en scène comme méthode propre à une réponse quantitative et à un prédictor quantitatif : ajuster une droite, interpréter la pente, tester si la pente est nulle. Agresti commence par les nuages de points, si bien que vous voyez la relation avant de la modéliser. Le coefficient de corrélation est réinterprété comme une pente standardisée, reliant les outils descriptifs antérieurs à la machine inférentielle. Le chapitre s'achève sur la prédiction et ses limites honnêtes.

La pente d'une régression répond à une question précise : de combien la réponse change-t-elle, en moyenne, pour un accroissement d'une unité du prédictor, tout le reste étant tenu hors du tableau. Une forte corrélation n'établit jamais la causalité, et c'est en extrapolant au-delà de la plage observée de x que les prédictions vont mourir. La droite est un résumé, non une loi. Les graphiques de résidus sont le diagnostic qui vous dit si le modèle mérite foi.

À essayer: Ajustez une droite à deux variables quantitatives de votre choix, puis tracez les résidus ; une courbure y signifie que la droite droite vous ment.

La régression multiple : les autres variables maintenues constantes

Voici la méthode centrale du livre : la régression multiple, où plusieurs prédictors partagent un même modèle. Agresti montre comment une pente change de sens en présence d'autres variables — elle devient l'effet d'un prédictor les autres étant maintenus fixes. Des exemples de revenu prédit par l'éducation et l'expérience professionnelle révèlent comment une relation bivariable peut rétrécir, s'inverser ou disparaître dès qu'une troisième variable entre en jeu. C'est la machinerie derrière le « en contrôlant pour » de chaque article en sciences sociales.

Le coefficient de régression multiple est le cheval de bataille des sciences sociales observationnelles parce qu'il approche ce qu'une expérience montrerait — en ajustant les facteurs de confusion mesurés. Mais l'ajustement ne vaut que pour les variables que vous avez mesurées et incluses ; les variables omises gardent leur biais. La multicolinéarité peut rendre les coefficients individuels instables même quand le modèle prédit bien. Lire un tableau de régression est un art d'interprétation, non d'arithmétique.

À essayer: Trouvez un tableau de régression dans un article de presse ou un rapport, et repérez quel coefficient porte l'affirmation de titre et par rapport à quelles variables il a été ajusté.

Comparer plusieurs groupes : l'analyse de variance

Quand il faut comparer plus de deux groupes, Agresti introduit l'analyse de variance : on décompose la variabilité totale en une part inter-groupes et une part intra-groupes, puis on teste si la part inter-groupes dépasse ce que le hasard autorise. Un exemple comparant les réponses moyennes entre plusieurs catégories de traitement montre la statistique F à l'œuvre. La significativité du test n'est que le premier geste ; des comparaisons de suivi identifient quels groupes diffèrent.

L'ANOVA est une régression avec des prédicteurs catégoriels déguisée comme telle, et voir cette équivalence constitue l'unification discrète du livre. Un test F significatif dit que les moyennes des groupes ne sont pas toutes égales, pas lesquelles diffèrent — des comparaisons planifiées ou des intervalles de confiance pour les paires doivent donc suivre. La variabilité intra-groupes est la règle qui décide si les différences inter-groupes sont impressionnantes. Multiplier les tests par paires sans ajustement gonfle la probabilité de faux positifs.

À essayer: Si vous comparez plus de deux groupes, planifiez une procédure de suivi pour les différences par paires avant de lancer le test F global.

Combiner les méthodes : construction de modèles et diagnostics

Vers la fin, Agresti assemble la boîte à outils : sélection de modèles, termes d'interaction et diagnostics qui vérifient si un modèle de régression s'ajuste bien. Les exemples montrent l'interaction — quand l'effet d'un prédicteur dépend du niveau d'un autre — changer entièrement l'interprétation. L'analyse des résidus et la vérification des hypothèses sont traitées comme une hygiène de routine, non comme des extras facultatifs. Le résultat : un chercheur capable de construire un modèle puis de l'interroger.

Un modèle est une affirmation sur la façon dont le monde génère les données, et les diagnostics sont la manière de tester votre propre affirmation. Les termes d'interaction comptent parce que des effets moyennés peuvent induire en erreur quand les effets diffèrent selon les sous-groupes. La sélection automatique de variables ne peut remplacer la réflexion sur les variables qui appartiennent au modèle. Les meilleurs analystes passent plus de temps à vérifier les modèles qu'à les ajuster.

À essayer: Dans votre prochaine régression, ajoutez un terme d'interaction dont vous suspectez qu'il compte et comparez si l'interprétation des effets principaux change.

La régression logistique pour les résultats binaires

Le livre clôt son arc de modélisation avec la régression logistique, le sujet de prédilection d'Agresti : modéliser une réponse binaire — succès ou échec, a voté ou non — à l'aide de prédicteurs. Il introduit les cotes et les rapports de cotes, puis montre comment le modèle relie

les prédicteurs à la probabilité de succès. Un exemple prédisant un résultat binaire à partir de prédicteurs continus et catégoriels montre la portée de la méthode. Le chapitre renvoie aux tableaux de contingence des chapitres antérieurs, désormais avec le pouvoir d'ajuster pour d'autres variables.

La régression logistique étend les tableaux croisés de l'ère du khi-deux vers des modèles à nombreux prédicteurs, voilà pourquoi elle domine la recherche en sciences sociales sur les résultats binaires. Les rapports de cotes sont la mesure d'effet naturelle, mais ce ne sont pas des rapports de risques, et les confondre exagère les effets quand le risque de base est élevé. Le modèle opère sur l'échelle des log-cotes, si bien que les coefficients doivent être traduits avant de signifier quoi que ce soit au lecteur. C'est la porte d'entrée vers les travaux avancés d'Agresti sur les données catégorielles.

À essayer: Prenez un résultat binaire qui vous tient à cœur et calculez à la main le rapport de cotes entre deux groupes avant d'ajuster le moindre modèle.

Limites

Le livre s'appuie sur les méthodes fréquentistes classiques ; les lecteurs en quête d'approches bayésiennes ou de techniques modernes de rééchantillonnage devront se tourner vers d'autres sources. Certaines éditions traitent brièvement de sujets comme la régression logistique et les modèles multiniveaux, laissant l'analyse catégorielle avancée aux ouvrages de troisième cycle d'Agresti. Les exemples, tirés de la recherche en sciences sociales, peuvent sembler datés, et le rythme s'accélère nettement après les premiers chapitres — les étudiants peu à l'aise en algèbre risquent de trouver le contenu sur la régression exigeant sans entraînement supplémentaire.

En une phrase

Quantifiez votre incertitude avant toute affirmation : décrivez les données, estimez avec des intervalles, et traitez chaque p-value comme une question plutôt que comme un verdict.

Pour qui

Les étudiants en sciences sociales qui suivent leur premier véritable cours de statistique, et les praticiens qui veulent une référence rigoureuse mais lisible sur la régression, l'ANOVA et l'analyse des données catégorielles. Quiconque a mémorisé des formules sans comprendre pourquoi elles fonctionnent trouvera dans ces explications connectées de quoi y voir clair.