

تاریخچه توجه به ملاحظات اخلاقی در انجام کارهای آماری

علی دولتی

گروه آمار، دانشکده علوم ریاضی، دانشگاه یزد
adolati@yazd.ac.ir

نامه علوم پایه شماره ۱۶، پاییز و زمستان ۱۴۰۴

چکیده

آمار در طول تاریخ، هم‌زمان به‌عنوان مجموعه‌ای از فنون کمی و حوزه‌ای از دانش با دلالت‌های اجتماعی و اخلاقی عمیق ایفای نقش کرده است. مقاله حاضر با رویکردی تاریخی به کاوش در سیر تحول توجه به ملاحظات اخلاقی در انجام کارهای آماری می‌پردازد. روایت مقاله از ریشه‌های تاریخی آمار آغاز می‌شود: از سرشماری‌های دوران باستان و سنت علم دولت در آلمان تا حساب سیاسی در انگلستان و پیوند آمار با جنبش اصلاح نژادی. در ادامه، با عبور از جنگ جهانی دوم و آزمایش‌های غیرانسانی نازی‌ها، منشور نورنبرگ به‌عنوان نخستین بیانیه بین‌المللی اخلاق در پژوهش معرفی می‌گردد. رسوایی مطالعه توسکگی که افشای آن به تحولی اساسی در اخلاق پژوهش انجامید، تحلیل می‌شود و سه اصل محوری گزارش بلومونت (احترام به اشخاص، سودرسانی و عدالت) بررسی می‌گردند. مقاله، فرایند نهادینه‌سازی اخلاق در آمار را دنبال کرده و تدوین نخستین راهنمای اخلاقی در انجمن آمار آمریکا و مؤسسه بین‌المللی آمار را مرور می‌کند. در ادامه، بحران تکرارپذیری و بازتولیدپذیری نتایج در پژوهش‌های علوم تجربی و نقش سوءاستفاده از روش‌های آماری واکاوی شده و بیانیه تاریخی انجمن آمار آمریکا درباره معناداری آماری تفسیر می‌گردد. مقاله با ورود به عصر داده‌های بزرگ، سه چالش اساسی بحران اخذ رضایت آگاهانه، سوگیری الگوریتمی و عدالت آماری و بازتعریف مسئولیت آماردان را بررسی می‌کند. در پایان، مقاله با نگاهی به وضعیت ایران، به بررسی جای خالی گفت‌وگوی تاریخی - نقادانه درباره اخلاق آماری پرداخته و بر ضرورت بازاندیشی در نسبت آمار با مسئولیت اجتماعی تأکید می‌نماید.

کلیدواژه‌گان: اخلاق آماری، تاریخ آمار، مطالعه توسکگی، بحران تکرارپذیری، عدالت الگوریتمی

مقدمه

اخلاق و آمار تمایز قائل شد. سطح نخست، «اخلاق در آمار»^۱ است که به رفتار حرفه‌ای آماردانان، صداقت در تحلیل، پرهیز از دست‌کاری نتایج و پایبندی به استانداردهای روش‌شناختی اشاره دارد. این سطح در دهه‌های اخیر و به‌ویژه با بحران تکرارپذیری و بازتولیدپذیری نتایج در پژوهش‌های علوم تجربی، توجه گسترده‌ای یافته است [۲۲]. اما سطح دوم، «اخلاق آمار»^۲ است که به این پرسش می‌پردازد: خود روش‌های آماری و نحوه کاربست آن‌ها در ساختارهای قدرت، چگونه می‌تواند به بازتولید نابرابری‌ها، توجیه بی‌عدالتی‌ها یا پنهان‌سازی آسیب‌های اجتماعی بینجامد؟ این سطح ریشه در تاریخ‌های کهن‌تری دارد و به دوران شکل‌گیری آمار به‌عنوان «علم حکومت‌داری» بازمی‌گردد [۴۰].

مقاله از هفت بخش اصلی تشکیل شده است. بخش نخست به ریشه‌های تاریخی آمار از دوران باستان تا سده نوزدهم می‌پردازد؛ از سرشماری‌های مصر، ایران، چین و روم باستان گرفته تا سنت آمار دانشگاهی در آلمان و حساب سیاسی در انگلستان. در ادامه،

آمار امروزه در بسیاری از حوزه‌های علمی و تصمیم‌گیری‌های اجتماعی چنان جایگاهی یافته که چه‌بسا بتوان آن را «زبان علم» یا «ابزار شناخت» نامید. این تصور که آمار صرفاً مجموعه‌ای از فنون کمی برای خلاصه‌سازی و تحلیل داده‌هاست، ریشه در سنت دیرپایی دارد که آماردان را ناظر خنثی و روش‌های آماری را ابزاری فارغ از ارزش می‌داند. اما تاریخ نشان می‌دهد که آمار در هر مرحله از رشد و کاربرد خود، با پرسش‌های اخلاقی اساسی گره‌خورده است؛ پرسش‌هایی که گاه در پس‌زمینه اطمینان از روش‌شناسی درست پنهان مانده‌اند و گاه با فجایع اخلاقی به سطح وجدان عمومی راه یافته‌اند [۵۴، ۷۴].

پرسش محوری این مقاله آن است که آگاهی از غیرفنی بودن صرف آمار چگونه پدید آمد؛ آگاهی از اینکه هر گام در پژوهش، از طرح تا تحلیل و گزارش، با ارزش‌ها، سوگیری‌ها و پیامدهای اخلاقی عجین است. این روایت، لایه‌های تاریخ پیچیده پیوند آمار و اخلاق را می‌گشاید. برای فهم این تحول، باید بین دو سطح از مسئله

دفتر آمار و جریده (فهرست سپاهیان) جزئی از دیوان حکومتی و دستگاه سلطنتی بوده است و این سنت تا دوره‌های بعدی نیز استمرار یافته است [۱۶]. اولین نشانه‌های عینی و تاریخی ظهور مفهوم دیوان در سلسله‌های پادشاهی ایرانی، مربوط به دوران داریوش اول هخامنشی است. در دوره سلوکیان و اشکانیان نظام حکومت‌داری هخامنشی ادامه یافت و به دوره ساسانیان منتقل شد. ساسانیان توجه بیشتری به آمار داشتند و امور مالی، کشاورزی و صنعتی و بازرگانی خود را بر اساس آمارهای جمع‌آوری شده به انجام می‌رساندند. امویان با کمک ایرانیان، برای اقتباس نظام دیوانی ساسانی تلاش زیادی کردند. در نخستین حکومت‌های ایرانی که بعد از اسلام تشکیل شدند (طاهریان، صفاریان، آل بویه و سامانیان)، نظام دیوانی دوره ساسانی به صورت دستگاهی منظم برای اداره امور کشور، مجدد ایجاد شد. در دوره سلجوقیان، نظریه‌پردازی خواجه نظام‌الملک در زمینه دیوان‌سالاری، زمینه‌ساز شکل‌گیری اصول و نظم اداری برای اداره کشور گردید. نظام دیوانی به صورت دولت فراگیر بعد از سلسله خوارزمشاهی و استیلای مغولان عملاً تعطیل شد. با تشکیل حکومت مرکزی در دوره صفوی، بعد از سه قرن شکلی جدید از دستگاه دیوانی ایجاد شد. در دیوان‌ها، مشاغل و منصب‌هایی وجود داشت که کار آنها در ارتباط با ثبت اطلاعات و آمار بود. به عنوان مثال لشکر شماری و حضور و غیاب سپاهیان، آمادگی و گسیل سربازان به جنگ، شمار کشتگان و اسیر شوندگان و تعداد جنگ‌افزارها از وظایف دیوان عرض بوده که سابقه آن به ایران باستان برمی‌گردد [۱۷]. در شاهنامه بارها از دیوان عرض یاد شده است. در داستان پادشاهی هرمز (پسر انوشیروان) صحنه‌ای از گفتگوی موبد و شاه برای تصمیم‌گیری در مورد آغاز یک نبرد به تصویر کشیده شده است:

به موبد چنین گفت جوینده راه
که اکنون چه سازیم با ساوه شاه
بدو گفت موبد که لشکر بساز
که خسرو به لشکر بود سرفراز
که چند است مردم که آید به کار
عرض را بخوان تا بیارد شمار
بیامد بیاورد بی‌مر سپاه
عرض با جریده به نزدیک شاه
پیاده بسی در میان سوار
شمار سپاه آمدش صدهزار
بدو گفت موبد که با ساوه شاه
سزد گر نشوریم با این سپاه

در این ابیات، «عرض» که مسئول شمارش لشکر است و «جریده» (دفتر یا فهرست آماری سپاه) نقشی تعیین‌کننده در تصمیم‌گیری جنگی ایفا می‌کنند [۱۷]. شاه پس از آگاهی از شمار سپاه (صدهزار نفر) با مشورت موبد، در مورد نبرد تصمیم می‌گیرد. این روایت

با بررسی آرای اندیشمندان این دوره، زمینه پیوند آمار با جنبش اصلاح نژادی در سده نوزدهم فراهم می‌شود و نقش پیش‌گامان آمار در این پیوند واکاوی می‌گردد. بخش دوم به جنگ جهانی دوم و آزمون‌های غیرانسانی نازی‌ها در دوران جنگ اختصاص دارد و از منشور نورنبرگ به عنوان نخستین بیانیه بین‌المللی اخلاق در پژوهش یاد می‌کند. بخش سوم به رسوایی مطالعه توسکگی می‌پردازد که افشای آن سرآغاز بیداری وجدان عمومی و ایجاد تحولی اساسی در اخلاق پژوهش بود؛ در ادامه، سه اصل کلیدی اخلاق پژوهش (احترام به اشخاص، سودرسانی و عدالت) و تأسیس نهادهای نظارتی تحلیل می‌شوند. بخش چهارم به فرایند نهادینه‌سازی اخلاق در آمار می‌پردازد و ضمن مرور پیشینه تدوین نخستین استانداردهای اخلاقی در انجمن‌های حرفه‌ای آمار، چالش‌های اخلاقی آمار دولتی و صنعتی را نیز بررسی می‌کند. بخش پنجم بحران تکرارپذیری و بازتولیدپذیری نتایج در علوم تجربی و نقش سوءاستفاده از روش‌های آماری را واکاوی می‌کند. بخش ششم به چالش‌های اخلاقی عصر داده‌های بزرگ و یادگیری ماشین می‌پردازد و مباحثی مانند بحران اخذ رضایت آگاهانه در پژوهش‌های کلان‌مقیاس، سوگیری الگوریتمی و عدالت آماری، و بازتعریف مسئولیت آماردان در زنجیره ارزش داده را بررسی می‌کند. بخش هفتم با نگاهی به وضعیت ایران، جای خالی گفت‌وگوی تاریخی - نقادانه درباره اخلاق آماری در آموزش و پژوهش را بررسی می‌کند و بر ضرورت بازاندیشی در نسبت آمار با مسئولیت اجتماعی در بستر تاریخی و فرهنگی ایران تأکید می‌نماید. مقاله با نتیجه‌گیری و جمع‌بندی درس‌های تاریخی و چشم‌انداز آینده اخلاق آماری به پایان می‌رسد.

۲. ریشه‌های تاریخی آمار

برای درک چرایی تأخیر در شکل‌گیری آگاهی اخلاقی در جامعه آماردانان، ناگزیر باید به سرچشمه‌های تاریخی این دانش بازگشت. آمار در آغاز راه، نه در دانشگاه‌ها و نه در متون صرفاً علمی، بلکه در دفاتر دولتی و به عنوان ابزاری برای حکومت‌داری و اداره سرزمین متولد شد. واژه Statistics (آمار) ریشه در واژه آلمانی Statistik و لاتین Status دارد که به معنای «دولت» یا «وضعیت سیاسی» است. این تبار دولتی - اداری، از همان آغاز آمار را در جایگاهی قرارداد که همواره در خدمت ساختارهای قدرت بود، بی‌آنکه پرسش از مشروعیت یا اخلاقی بودن آن ساختارها بخشی از وظایف آماردان محسوب شود [۵۴، ۴۰].

۲-۱ سرشماری در دوران باستان

استفاده از آمار به شکل سرشماری و برآوردهای مالیاتی از بدو پیدایش تمدن‌های باستانی (ایران، مصر، چین و روم باستان) ملازم با حکمرانی مؤثر و اساس اداره حکومت به ویژه در امر لشکررداری و لشکرکشی بوده است. به عنوان مثال، در ایران باستان، نگهداری

۲-۲-۲. حساب سیاسی در انگلستان

«حساب سیاسی»^۸ در انگلستان پدید آمد، در اواخر سده هفدهم شکوفا شد و در طول سده هجدهم نمایندگانی نیز در قاره اروپا یافت. در حالیکه آمار دانشگاهی عمدتاً ماهیتی توصیفی داشت، حساب سیاسی بر جنبه تحلیلی آمار تأکید می‌کرد. این جریان که در پی دقت بود، بر داده‌های عددی تکیه داشت و روابط بین آمارهای جمعیتی و اجتماعی و پدیده‌هایی مانند رشد جمعیت، تولدها، مرگ‌ومیر و ازدواج‌ها را مورد مطالعه قرار می‌داد [۴۰، ۷۴]. جان گرانت^۹ (۱۶۷۴-۱۶۲۰)، بازرگان و عضو شورای شهر لندن، معمولاً پدر جمعیت‌شناسی شناخته می‌شود [۱۶]. او در سال ۱۶۶۲ کتاب «مشاهدات طبیعی و سیاسی درباره صورت‌های مرگ‌ومیر» را منتشر کرد و در آن کوشید نخستین جدول مرگ‌ومیر را تدوین کند. گرانت از ثبت‌های هفتگی تولد و مرگ در لندن که از سال ۱۶۰۳ منتشر می‌شدند، استفاده کرد. این صورت‌های مرگ‌ومیر در اصل برای پیگیری همه‌گیری طاعون به کار می‌رفتند. با وجود آنکه این داده‌ها شامل اطلاعات سنی نبودند، گرانت به شیوه‌ای ابتکاری جمعیت کل لندن، توزیع سنی آن و حتی جدول مرگ‌ومیر را برآورد کرد [۴۰، ۷۴]. ویلیام پتی^{۱۰} (۱۶۸۷-۱۶۲۳) پژوهش‌های دوست خود جان گرانت را با تأکید بیشتر بر مسائل اقتصادی ادامه داد و اصطلاح حساب سیاسی را رواج داد و بنیان علمی تازه‌ای پایه‌ریزی کرد [۴۰، ۷۴]. زمینه‌سازی برای تأسیس دفاتر آماری مدرن در اروپا، نخستین‌بار توسط ویلیام پتی در انگلستان مطرح شد. چند سال بعد، گوتفرد لایبنیتس^{۱۱} (۱۷۱۶-۱۶۴۶) که نقشی برجسته در سیاست اروپای زمان خود ایفا می‌کرد، ایجاد تشکیلاتی با عنوان دفتر آمار را به امپراتوری پروس توصیه کرد. او بر این باور بود که دولت باید یک دفتر مرکزی آمار داشته باشد تا از قدرت خودآگاهی یابد [۱۶]. پس از این دوره، با انسجام یافتن دولت‌های مرکزی اروپا، دفاتر آماری به تدریج در این کشورها تأسیس شدند [۴۰].

۲-۲-۳. تکوین نظریه احتمال

توسعه نظریه احتمال در سده‌های هفدهم و هجدهم موجب جهشی عظیم در روش‌شناسی آماری شد. نخستین رویکردها به نظریه احتمال که با تحلیل بازی‌های شانسی پیوند داشتند، در سده شانزدهم پدیدار شدند. ریشه‌های نظریه احتمال را می‌توان در تحلیل بازی‌های شانسی جستجو کرد [۴۰]، اما شکل‌گیری آن به‌عنوان یک حوزه علمی منسجم در قرن هفدهم و با آثار بلز پاسکال و پیر دو فرما تحقق یافت. مسئله تقسیم عادلانه امتیازها در بازی ناتمام، نخستین مسئله‌ای بود که به صورت‌بندی دقیق مفهوم امید ریاضی و احتمال انجامید [۱۶]. یاکوب برنولی^{۱۲} (۱۷۱۳) با اثبات قضیه حد مرکزی

نشان می‌دهد که در ایران باستان، آمار افزون بر ابزار ثبت، مبنایی برای کنش سیاسی و نظامی نیز بوده است.

۲-۲-۲. از علم دولت و حساب سیاسی تا آمار مدرن

گذار به دوران مدرن با ظهور دولت‌های متمرکز اروپایی و نظام‌های اقتصادی نوین، ضرورت توجه به آمار برنامه‌ریزی‌شده را رقم زد. در قرن هجدهم، دو رویکرد نظری متفاوت در مورد آمار در اروپا شکل گرفت: در آلمان آمار دانشگاهی با نگاهی کل‌نگر و کیفی، و در انگلستان حساب سیاسی با رویکردی کمی و تحلیلی [۴۰].

۲-۲-۱. آمار دانشگاهی در آلمان

هرمان کونرینگ^۱ (۱۶۸۱-۱۶۰۶) در دانشگاه هلمشتت^۲ نخستین سخنرانی آمار را با عنوان علوم سیاسی ارائه داد. مهم‌ترین نماینده این سنت، گوتفرد آخنوال^۳ (۱۷۷۲-۱۷۱۹) در دانشگاه گوتینگن^۴ بود که واژه آمار را به‌صورت منسجم در معنای دانش امور دولت یا آمار حکومتی به کار برد [۴۰]. پس از او، آگوست لودویگ فون شلوتسر^۵ (۱۸۰۹-۱۷۳۵) جانشین وی در دانشگاه گوتینگن شد. او آمار را تاریخ ایستا تعریف می‌کرد و خود تاریخ را نوعی آمار پویا و پیوسته می‌دانست [۱۶]. از مهم‌ترین آثار او می‌توان به مکاتباتی عمدتاً با محتوای تاریخی و سیاسی و نیز انتشار نشریه‌ای که افزون بر ارائه داده‌ها و اطلاعات آماری، به اهمیت سیاسی برخی واقعیت‌های آماری نیز می‌پرداخت. همین رویکرد انتقادی و سیاسی سبب شد که نشریه او برای چند سال ممنوع شود؛ رخدادی که نشان می‌دهد آمار، حتی در سده هجدهم، دانشی صرفاً فنی و خنثی تلقی نمی‌شد، بلکه ظرفیتی انتقادی و حتی بالقوه انقلابی داشت. مناقشه مکتبی میان آماردانان دانشگاهی برای مدت طولانی ادامه یافت. سرانجام، دیدگاه موسوم به «آماردانان جدولی»^۶ (که صرفاً به تولید و ارائه جداول عددی اکتفا می‌کردند) غلبه پیدا کرد. زیرا جداول امکان سازمان‌دهی منسجم و نظام‌مند مواد آماری را فراهم می‌کردند. در این رویکرد، نمایش آماری نه صرفاً ابزاری کمکی، بلکه بخشی اساسی از سامان‌دهی دانش آماری به شمار می‌رفت. آگوست فریدریش ویلهلم کرومه^۷ (۱۸۳۳-۱۷۵۳) از جمله این آماردانان جدولی بود که نخستین نمودارهای میله‌ای و مساحتی را توسعه داد. یکی از شاخه‌های آمار دانشگاهی «آمار موضوعی و مفهومی» بود. پژوهش‌های آماری درباره خودکشی، جایگاه زنان در زندگی فرهنگی و موضوعات مشابه در زمره «آمار موضوعی» قرار می‌گرفتند. «آمار مفهومی» بر بنیان‌های مفهومی گردآوری داده‌ها، تفسیر زبانی آمار توصیفی و آمار اقتصادی تأکید داشت و نفوذ آن فراتر از محیط دانشگاهی نیز گسترش یافت [۴۰].

1. Hermann Conring
2. Helmstedt
3. Gottfried Achenwall
4. Göttingen
5. August Ludwig von Schlözer
6. Tabular statisticians

7. August Friedrich Wilhelm Crome
8. Political Arithmetic
9. John Graunt
10. William Petty
11. Gottfried Leibniz
12. Jakob Bernoulli

حساب‌های ملی ظاهر شد و بخشی نیز در جمعیت‌شناسی جذب گردید. در آستانه انقلاب صنعتی، دو رویکرد متفاوت به پویایی جمعیت شکل گرفت:

رویکرد الهیاتی: یوهان پیتر سوس میلش^{۱۱} (۱۷۶۷-۱۷۰۷) در کتاب خود «نظم الهی در تغییرات نوع بشر» کوشید الگوهای جمعیتی را در چارچوب مشیت الهی تبیین کند. او معتقد بود پدیده‌های جمعیتی مانند تولد، ازدواج و مرگ تحت مشیت الهی هستند و نظم‌های آماری در این پدیده‌ها، خود نشانه‌ای از کار خداوند است [۴۳].

رویکرد اقتصادی: چند دهه بعد، توماس رابرت مالتوس^{۱۲} (۱۸۳۴-۱۷۶۶) در اثر مشهور خود با عنوان «رساله‌ای در باب اصول جمعیت»، دیدگاهی بدبینانه‌تر ارائه کرد. به باور او، جمعیت تمایل دارد به‌صورت هندسی رشد کند، درحالی‌که تولید غذا تنها رشد حسابی دارد؛ بنابراین، فشار جمعیتی ناگزیر از طریق جنگ، قحطی، بیماری و فقر مهار خواهد شد [۴۴]. نظریه مالتوس تأثیر عمیقی بر اقتصاد سیاسی، جمعیت‌شناسی و حتی زیست‌شناسی تکاملی گذاشت.

در سده نوزدهم، هم زمان با شکل‌گیری دولت‌های مدرن و گسترش اقتصاد بازار، آمار رسمی و سرشماری‌های ملی به‌سرعت توسعه یافتند. آدولف کوتله^{۱۳} (۱۷۹۶-۱۸۷۴) یکی از چهره‌های محوری این تحول بود. او در سازمان‌دهی نخستین سرشماری مدرن بلژیک (۱۸۴۶) نقش داشت و بعدها در تأسیس کنگره بین‌المللی آمار (۱۸۵۳) مشارکت کرد. کوتله بر این باور بود که جرم، مالیاتی است که جامعه باید بپردازد. او بر پایه میانگین‌های جمعیتی، مفهوم «انسان متوسط»^{۱۴} را بنا نهاد؛ شخصی که بر اساس میانگین ویژگی‌های یک جمعیت تعریف می‌شود و به‌عنوان انسان متوسط شناخته می‌شود. هرچند چنین فردی در واقعیت وجود خارجی ندارد، اما کوتله آن را راهنمایی برای تصمیم‌گیری‌های سیاسی و اجتماعی تلقی می‌کرد. او حتی این انسان متوسط را نوعی آرمان زیبایی‌شناختی می‌دانست. اندیشه کوتله درباره وجود نظم‌های آماری در پدیده‌های اجتماعی، تأثیر گسترده‌ای بر علوم اجتماعی و زیستی گذاشت. با این حال، در اواخر قرن نوزدهم و اوایل قرن بیستم، برخی جریان‌ها این ایده را به‌گونه‌ای تفسیر کردند که انحراف از میانگین را نشانه نقص یا انحراف اجتماعی تلقی می‌کرد. چنین برداشت‌هایی بعدها در برخی سیاست‌های اصلاح‌نژادی و عقیم‌سازی اجباری مورد استفاده قرار گرفت و یکی از بحث‌برانگیزترین فصل‌های تاریخ آمار و علوم اجتماعی را رقم زد [۴۰].

اولیه (آنچه امروز قانون اعداد بزرگ نامیده می‌شود)، رابطه میان احتمال نظری و فراوانی نسبی تجربی را تثبیت کرد و تمایز میان دیدگاه‌های پیشینی و پسینی در احتمال را برجسته ساخت [۴۰]. برنولی خود این نتیجه را قضیه طلایی نامید و آن را حاصل بیست سال اندیشه‌ورزی خویش معرفی کرد. در قرن هجدهم، این نظریه با آثار آبراهام دموآر^۱، توماس بیژ^۲ و پییر سیمون لاپلاس^۳ به یک دستگاه ریاضی منسجم تبدیل شد. دموآر توزیع نرمال را به‌عنوان حد توزیع دوجمله‌ای معرفی کرد و بدین ترتیب نخستین صورت از قضیه حد مرکزی را ارائه داد. توماس بیژ بنیان استنباط آماری مبتنی بر احتمال شرطی را فراهم ساخت، و لاپلاس صورت‌بندی کلاسیک احتمال را نظام‌مند کرد و نظریه تحلیلی احتمال را به اوج خود رساند [۱۶]. در همین دوره، کارل فردریش گاوس^۴ توزیع نرمال را در تحلیل خطاهای مشاهدات نجومی به کار گرفت و روش کمترین مربعات را به عنوان ابزار استاندارد برآورد توسعه داد. در قرن بیستم، آندری کولموگوروف^۵ (۱۹۳۳) با ارائه دستگاه اصل موضوعی، بنیان ریاضی نظریه احتمال را به طور مستقل از تفسیرهای فلسفی تثبیت نمود [۱۶].

۲-۲-۴. توسعه آمار مدرن

در قرن نوزدهم و اوایل قرن بیستم، فرانسیس گالتون^۶، کارل پیرسون^۷ و رونالد فیشر^۸ ابزارهای اصلی آمار استنباطی از جمله رگرسیون، همبستگی، آزمون فرض و تحلیل واریانس را توسعه دادند. گالتون مفاهیم همبستگی و رگرسیون را به درون آمار آورد، پیرسون ضریب همبستگی را بنیان نهاد، و فیشر روش‌های طراحی آزمایش، تحلیل واریانس و برآورد ماکسیمم درست‌نمایی را ابداع کرد [۴۰]. هم زمان ریاضی‌دانان روسی در شکل‌گیری نظریه حدی و فرایندهای تصادفی نقش اساسی ایفا کردند. در ادامه، جرجی نیمان^۹ و ایگون پیرسون^{۱۰} (پسر کارل پیرسون) چارچوب نظریه آزمون فرضیه‌ها را به‌صورت نظام‌مند تدوین کردند [۱۶]. در آستانه و پس از جنگ جهانی دوم، مجموعه‌ای از حوزه‌های میان‌رشته‌ای مرتبط با آمار مانند اقتصادسنجی، زیست‌سنجی، انسان‌سنجی، روان‌سنجی، جامعه‌سنجی و پژوهش عملیاتی شکل گرفتند. این تحولات نشان‌دهنده گذار آمار از یک ابزار اداری و دولتی به یک زبان تحلیلی عمومی برای مدل‌سازی پدیده‌های پیچیده در علوم طبیعی، اجتماعی و فنی است [۷۴، ۴۰].

۲-۳. آمار و جنبش اصلاح نژادی

بخش بزرگی از دستاوردهای حساب سیاسی بعدها در نظام

1. Abraham de Moivre
2. Thomas Bayes
3. Pierre-Simon Laplace
4. Carl Friedrich Gauss
5. Andrey Kolmogorov
6. Francis Galton
7. Karl Pearson

8. Ronald Fisher
9. Jerzy Neyman
10. Egon Pearson
11. Johann Peter Süssmilch
12. Thomas Robert Malthus
13. Adolphe Quetelet
14. Average man

باشد. پرسش از انتخاب مسئله‌ها، تعریف متغیرها، طبقه‌بندی جمعیت‌ها و کاربرد نتایج، همگی پرسش‌هایی اخلاقی هستند که در دوران پیشانهادهی آمار، به‌ندرت موضوع تأمل حرفه‌ای قرار می‌گرفتند. یکی از دلایل مهم غیاب آگاهی اخلاقی در این دوره، نبود نهادهای حرفه‌ای مستقل با استانداردهای اخلاقی الزام‌آور بود. نخستین انجمن‌های آماری (مانند انجمن آمار لندن، تأسیس ۱۸۳۴) بیشتر بر ارتقای روش‌های فنی و گسترش کاربردهای دولتی آمار متمرکز بودند تا بر تعیین حدود مسئولیت اخلاقی آماردان. در ایالات متحده نیز انجمن آمار آمریکا که در سال ۱۸۳۹ تأسیس شد، تا دهه‌های میانی قرن بیستم هیچ شیوه‌نامه اخلاقی مدونی نداشت [۲۱]. این خلأ نهادی به این معنا بود که تصمیم درباره‌ی این که چه پژوهش‌هایی از نظر اخلاقی مجاز هستند عمدتاً به وجدان فردی آماردانان واگذار شده بود، درحالی که فشارهای نهادی و منافع دولتی اغلب در جهت مخالف اخلاق عمل می‌کردند. این وضعیت تا میانه‌ی قرن بیستم و وقوع فجایع انسانی جنگ جهانی دوم ادامه یافت. فجایی که نشان دادند اتکای صرف به وجدان فردی و روش‌شناسی صحیح، برای جلوگیری از سوءاستفاده‌های نظام‌مند از آمار و داده‌های انسانی کافی نیست.

بخش بعدی این مقاله به بررسی این رویداد مهم می‌پردازد: دوره‌ای که جامعه جهانی پس از محاکمات دادگاه نورنبرگ، نخستین گام‌های جدی را برای نهادهای سازی اخلاق در پژوهش برداشت و آماردانان نیز ناگزیر به بازاندیشی در مسئولیت حرفه‌ای خود شدند.

۳. تولد اخلاق پژوهش

سده نوزدهم و آغاز سده بیستم، دوران پیوند آمار با دولت‌ها و طرح‌های اصلاح نژادی بود، اما میانه قرن بیستم، سرآغاز تحولی اساسی در اخلاق پژوهش به شمار می‌رود. جنگ جهانی دوم و جنایاتی که در پناه آن رخ داد، پندار بی‌طرفی علم را برای همیشه زدود. در این برهه، جامعه جهانی برای نخستین بار با این پرسش روبه‌رو شد که آیا هر پژوهشی، صرفاً به دلیل دقت فنی و روش‌شناختی، به طور خودکار موجه است؟ پاسخ به این پرسش، به شکل‌گیری اسناد و نهادهایی انجامید که امروز پایه‌های اخلاق پژوهش را تشکیل می‌دهند [۲۰].

۳-۱. پیامدهای نظریه اصلاح نژادی

جنبش اصلاح نژادی که نظریه آن را گالتون بنا نهاده بود، در اوایل قرن بیستم به اوج نفوذ خود رسید و در ایالات متحده به سرعت نهادینه شد. دفاتر ثبت اصلاح نژادی به مراکز اصلی گردآوری داده‌های خانوادگی و ترویج سیاست‌های نژادی تبدیل شدند [۳۸]. نخستین قانون عقیم‌سازی اجباری در جهان در سال ۱۹۰۷ در ایالت ایندیانا^۳ آمریکا به تصویب رسید. سپس در سال ۱۹۲۲،

جنبش اصلاح نژادی در دهه‌های پایانی قرن نوزدهم، ابتدا در بریتانیا و سپس در ایالات متحده و دیگر کشورهای اروپایی گسترش یافت. فرانسیس گالتون، پسر عموی چارلز داروین که معمولاً پدر آمار زیستی شناخته می‌شود، مفاهیمی چون همبستگی و رگرسیون را وارد آمار کرد و نقش مهمی در توسعه روش‌های آماری داشت. با این حال، گالتون بنیان‌گذار نظریه اصلاح نژادی نیز بود. او در کتاب نبوغ موروثی (۱۸۶۹)، با استفاده از داده‌های خانوادگی نخبگان بریتانیا، استدلال کرد که توانایی‌های ذهنی و اخلاقی عمدتاً موروثی‌اند و بنابراین جامعه باید از طریق سیاست‌های جمعیتی، کیفیت زیستی نسل‌های آینده را بهبود بخشد [۵۱]. گالتون واژه «اوزنیک»^۱ را به معنای «به زاد» یا «خوب زاد» ابداع کرد و آن را علم «بهبود نژاد» یا «اصلاح نژادی» نامید. کارل پیرسون، شاگرد و جانشین گالتون در کالج دانشگاهی لندن، این برنامه را نظام‌مندتر ساخت. او نخستین کرسی دانشگاهی آمار را بنیان گذاشت و مجله بیومتریکا را تأسیس کرد که همچنان از نشریات معتبر آمار زیستی است. پیرسون از کاربرد روش‌های آماری برای مطالعه تفاوت‌های انسانی دفاع می‌کرد و در بسیاری از آثارش از سیاست‌های اصلاح نژادی حمایت می‌نمود [۷۲، ۶۳]. رونالد فیشر نیز، با وجود نقش اساسی خود در بنیان‌گذاری آمار مدرن (از جمله تحلیل واریانس، طراحی آزمایش‌ها و نظریه برآورد) از حامیان جنبش اصلاح نژادی بود. او در دهه‌های ۱۹۲۰ و ۱۹۳۰ در نوشته‌های مختلف از سیاست‌های تشویق باروری طبقات مرفه و محدودسازی باروری گروه‌های فقیر حمایت کرد و با انجمن اصلاح نژادی بریتانیا همکاری داشت [۴۸، ۳۰]. امروزه بخش قابل توجهی از این دیدگاه‌ها به دلیل اتکا بر پیش‌فرض‌های نژادگرایانه و طبقاتی، از منظر علمی و اخلاقی مردود شمرده می‌شوند.

۲-۴. فقدان نهادهای اخلاقی

نکته حائز اهمیت برای تاریخ اخلاق آماری این است که گالتون، پیرسون و فیشر، هیچ یک سوءرفتار علمی به معنای متعارف (مانند جعل داده یا دست‌کاری نتایج) مرتکب نشدند. آن‌ها روش‌های آماری دقیق و پیشرفته‌ای به کار می‌گرفتند و در چارچوب استانداردهای علمی زمان خود عمل می‌کردند. با این حال، پژوهش‌های آن‌ها در خدمت پروژه‌ای قرار داشت که امروزه آشکارا غیراخلاقی تشخیص داده می‌شود: پروژه‌ای که به توجیه نابرابری‌های نژادی و طبقاتی، ترویج سیاست‌های «عقیم‌سازی اجباری»^۲ و در نهایت (در موارد افراطی‌تر مانند آلمان نازی) به مشروعیت‌بخشیدن به نسل‌کشی انجامید [۷۵]. این تجربه تاریخی نشان می‌دهد که اخلاق در آمار فراتر از صداقت فنی و پرهیز از جعل داده است. حتی پژوهشگری که صادقانه و با دقیق‌ترین روش‌ها کار می‌کند، ممکن است در چارچوبی فعالیت کند که خود آن چارچوب ناعادلانه و تبعیض‌آمیز

1. Eugenics
2. Sterilization

3. Indiana

به محاکمه پزشکان، به بررسی جنایات پزشکان و پژوهشگران آلمانی اختصاص داشت. در این محاکمه، ۲۰ پزشک و سه مقام اداری متهم شدند که در آزمایش‌های غیرانسانی مشارکت داشتند [۲۰]. نکته حائز اهمیت از منظر تاریخ اخلاق پژوهش این است که متهمان در دفاع از خود، به استدلالی رجوع کردند که تا پیش از آن در محافل علمی رایج بود: پژوهش‌های ما از نظر روش‌شناختی دقیق بوده و نتایج آن‌ها به نفع علم و بشریت است. آن‌ها مدعی بودند که ارزش علمی نتایج، می‌تواند فقدان رضایت آزمودنی‌ها و آسیب‌های وارد شده را توجیه کند. دادگاه نورنبرگ این استدلال را صریحاً رد کرد. در ۱۹ اوت ۱۹۴۷، قضات ده ماده را به‌عنوان اصول مجاز بودن آزمایش‌های پزشکی اعلام کردند که بعدها به «منشور نورنبرگ»^۵ معروف شد. این سند که نخستین بیانیه بین‌المللی در حوزه اخلاق پژوهش محسوب می‌شود، چند اصل اساسی را تعیین کرد که مهم‌ترین آن‌ها عبارت بودند از [۶۹]:

رضایت داوطلبانه و آگاهانه: آزمودنی باید از توانایی قانونی برای ابراز رضایت برخوردار باشد و پس از آگاهی کامل از ماهیت، مدت، اهداف، روش‌ها، خطرات و پیامدهای پژوهش، آزادانه رضایت خود را اعلام نماید.

نتایج مفید برای جامعه: پژوهش باید به‌گونه‌ای طراحی شود که نتایج آن به نفع جامعه باشد و با استفاده از روش‌های دیگر قابل‌دستیابی نباشد.

اجتناب از رنج و آسیب غیرضروری: پژوهش باید بر اساس یافته‌های پیشین طراحی گردد تا رنج و آسیب جسمی و روحی به حداقل ممکن کاهش یابد.

نسبت خطر به منفعت: خطرات ناشی از پژوهش نباید از اهمیت انسانی مسئله فراتر رود.

حق انصراف آزمودنی: آزمودنی باید بتواند در هر مرحله از پژوهش، از ادامه مشارکت انصراف دهد.

توقف پژوهش در صورت بروز خطر: پژوهشگر موظف است در صورت مشاهده هرگونه خطر برای آزمودنی، پژوهش را متوقف کند

۳-۳. مطالعه توسکگی

منشور اخلاقی نورنبرگ اگرچه نخستین بیانیه بین‌المللی اخلاق در پژوهش بود، اما «مطالعه توسکگی»^۶ بود که این اصول را از سطح اسناد آلمانی به عرصه نهادهای الزام‌آور فرود آورد و به‌عنوان چرخش‌گاهی تاریخی در این زمینه شناخته شد. این مطالعه که از ۱۹۳۲ تا ۱۹۷۲ به مدت چهل سال ادامه یافت، به یکی از بزرگ‌ترین رسوایی‌های اخلاقی در تاریخ علم آمریکا تبدیل شد و

هری لافلین^۱ قانون نمونه‌ای را تدوین کرد که بعدها در آلمان نازی الهام‌بخش قوانین مشابه گردید [۶۱]. پرونده معروف «باک علیه بل»^۲ (۱۹۲۷) در دیوان عالی آمریکا، عقیم‌سازی اجباری افراد دارای ناتوانی ذهنی را قانونی اعلام کرد. قاضی صادرکننده رأی، در جمله‌ای مشهور گفت: «سه نسل از احمق‌ها کافی است» [۳۳]. از سال ۱۹۰۷ تا دهه ۱۹۷۰، بیش از ۶۰۰۰۰ نفر در ۳۳ ایالت آمریکا تحت عقیم‌سازی اجباری قرار گرفتند [۶۲]. در آلمان نیز این ایده‌ها با استقبال گسترده‌ای روبه‌رو شد. نازی‌ها از همان ابتدای به‌قدرت‌رسیدن، قوانین اصلاح نژادی را با شدت بیشتری به اجرا درآوردند. قانون جلوگیری از تولد فرزندان مبتلا به بیماری‌های ارثی (۱۹۳۳) که به عقیم‌سازی اجباری بیش از ۴۰۰۰۰ نفر انجامید، با استناد به داده‌های آماری درباره هزینه اجتماعی بیماری‌های ارثی توجیه می‌شد [۷۵]. آماردانان آلمانی در گردآوری و تحلیل این داده‌ها نقش فعالی داشتند و بسیاری از آن‌ها خود از حامیان ایدئولوژی نازی بودند. قوانین نورنبرگ^۳ (۱۹۳۵) که مبنای حقوقی آزار و اذیت یهودیان را فراهم کرد، مستقیماً از مدل‌های نژادپرستانه آمریکایی الهام گرفته شده بود [۶۰]. برنامه موسوم به «تی چهار»^۴ (۱۹۴۱-۱۹۳۹) کشتن نظام‌مند نزدیک به ۷۰۰۰۰ معلول ذهنی و جسمی را سازماندهی کرد [۳۴].

۳-۲. پزشکی و آمار در خدمت نازیسم

رژیم نازی در آلمان (۱۹۳۳-۱۹۴۵) از علم به‌ویژه پزشکی و آمار به‌عنوان ابزاری برای مشروعیت‌بخشی به سیاست‌های نژادپرستانه خود استفاده می‌کرد. اما فراتر از قوانین اصلاح نژادی، آنچه تاریخ اخلاق پژوهش را برای همیشه دگرگون کرد، آزمایش‌های پزشکی غیرانسانی بود که در اردوگاه‌های کار اجباری بر روی زندانیان انجام می‌گرفت. این آزمایش‌ها که بدون رضایت آگاهانه و با تحمیل درد و رنج عظیم انجام می‌شدند، شامل آزمایش‌های غواصی در آب یخ (برای بررسی بقا در آب‌های سرد شمالی)، آزمایش‌های ارتفاع (برای شبیه‌سازی شرایط سقوط خلبانان در ارتفاع بالا)، آزمایش‌های مالاریا، فسفر، گاز خردل و بسیاری موارد دیگر بودند [۲۰]. در این آزمایش‌ها، آمار در طراحی و تحلیل نقش داشت، و گاهی به ابزاری برای کاهش هزینه انسانی پژوهش تبدیل می‌شد. برای نمونه، در آزمایش‌های غواصی در آب یخ، پژوهشگران با استفاده از روش‌های آماری به دنبال یافتن حداقل دمای لازم برای ایجاد هیپوترمی کشنده (کاهش دمای خطرناک بدن) بودند تا بتوانند تعداد قربانیان را بهینه‌سازی کنند [۲۸]. از منظر روش‌شناختی، این پژوهش‌ها گاه با استانداردهای فنی زمان خود تطبیق داشتند، اما از نظر اخلاقی، آشکارترین مصداق علم در خدمت جنایت بودند. پس از پایان جنگ، محاکمات دادگاه نورنبرگ برای رسیدگی به جنایات نازی‌ها برگزار شد. بخشی از این محاکمات، معروف

1. Harry Laughlin
2. Buck v. Bell
3. Nuremberg Laws

4. T4 Action
5. Nuremberg Code
6. Tuskegee study

شده توسط دیگران است. این نگاه تفکیک‌گرا بین طراحی اخلاقی و تحلیل فنی بعدها به چالش کشیده شد. سؤال اساسی این بود: آیا آماردانی که می‌داند داده‌ها از طریق فریب و با انکار درمان به دست آمده‌اند، می‌تواند صرفاً باتکیه بر دقت روش‌شناختی تحلیل خود، مسئولیت اخلاقی را از خود سلب کند؟ [۲۴]. افشای مطالعه توسسگی در ۱۹۷۲، موجی از اعتراضات عمومی را برانگیخت و کنگره آمریکا را به اقدام واداشت. در سال ۱۹۷۴، کنگره قانون پژوهش ملی را تصویب کرد که طی آن کمیسیون ملی برای حفاظت از آزمودنی‌های انسانی در پژوهش‌های زیست‌پزشکی و رفتاری تشکیل شد. این کمیسیون مأموریت یافت اصول اخلاقی را تدوین کند که زیربنای همه پژوهش‌های انسانی در ایالات متحده باشد [۲۷]. نتیجه چهار سال کار این کمیسیون، گزارش «بلومونت»^۳ بود که در ۱۹۷۹ منتشر شد. این گزارش سه اصل اساسی را معرفی کرد که تا امروز سنگ بنای اخلاق پژوهش در سراسر جهان محسوب می‌شوند:

۱. **احترام به اشخاص:** این اصل دو الزام اخلاقی را در بر می‌گیرد: نخست، به رسمیت‌شناختن خودمختاری افراد؛ دوم، حفاظت از کسانی که خودمختاری آن‌ها کاهش یافته است. ترجمه عملی این اصل در پژوهش، اخذ رضایت آگاهانه از مشارکت‌کنندگان است.

۲. **سودرسانی:** این اصل پژوهشگران را ملزم می‌کند که منافع بالقوه پژوهش را به حداکثر و آسیب‌های احتمالی را به حداقل برسانند. ترجمه عملی این اصل، ارزیابی دقیق نسبت خطر به منفعت، پیش از اجرای پژوهش است.

۳. **عدالت:** این اصل به توزیع عادلانه بارها و منافع پژوهش اشاره دارد. مطالعه توسسگی نمونه‌ای آشکار از نقض این اصل بود: بار پژوهش بر دوش جمعیتی فقیر و به حاشیه رانده شده بود، درحالی‌که منافع آن نصیب جامعه غالب می‌شد [۲۷].

نوآوری بزرگ گزارش بلومونت، صرفاً تدوین اصول اخلاقی نبود، بلکه نهادهای سازی آن‌ها از طریق سازوکارهای الزام‌آور بود. بر اساس قوانین مصوب پس از این گزارش، هر پژوهش انسانی که با بودجه فدرال انجام می‌شد، می‌بایست توسط هیئت‌های بازبینی نهادی مورد بررسی و تأیید قرار می‌گرفت. این هیئت‌ها که در دانشگاه‌ها و مراکز پژوهشی مستقر بودند، وظیفه داشتند اطمینان حاصل کنند که پژوهش‌های انسانی بر اساس سه اصل بلومونت طراحی و اجرا می‌شوند [۲۵]. برای آماردانان، این تحول به معنای ورود به عرصه تازه‌ای از مسئولیت‌های اخلاقی بود. دیگر نمی‌شد ادعا کرد که وظیفه آماردان تنها تحلیل صحیح داده‌ها است. آماردانان در جایگاه مشاوران آماری پژوهش‌های بالینی و رفتاری، ناگزیر به مشارکت در طراحی پژوهش‌ها و ارزیابی جنبه‌های اخلاقی آن‌ها شدند.

پیامدهای آن، تمامی حوزه‌های پژوهش‌های انسانی از جمله آمار را برای همیشه دگرگون کرد [۳۱، ۵۸]. مطالعه توسسگی توسط سرویس بهداشت عمومی ایالات متحده در شهر ماکون^۱، ایالت آلاباما^۲، در میان جمعیت فقیر و عمدتاً بی‌سواد سیاه‌پوستان آغاز شد. عنوان رسمی این پژوهش مطالعه سفلیس درمان نشده در مردان سیاه‌پوست بود. هدف پژوهش، بررسی روند طبیعی بیماری سفلیس در صورت عدم درمان بود [۵۸]. گروهی شامل ۶۰۰ مرد سیاه‌پوست در این مطالعه شرکت داشتند: ۳۹۹ مرد مبتلا به سفلیس (گروه آزمون) و ۲۰۱ مرد سالم (گروه کنترل). به این مردان گفته شده بود که برای درمان بیماری خونی بدخیم که اصطلاحی محلی برای طیفی از بیماری‌ها بود، تحت معاینه و درمان قرار می‌گیرند. در ازای مشارکت، آن‌ها از معاینه رایگان، وعده‌های غذایی گرم، و بیمه تشییع جنازه برخوردار می‌شدند [۳۱]. اما نکته فاحش غیراخلاقی این پژوهش آن بود که پس از شروع مطالعه، به شرکت‌کنندگان هیچ درمانی برای سفلیس ارائه نشد. حتی در سال ۱۹۴۳ که پنی‌سیلین به‌عنوان درمان مؤثر سفلیس شناخته شد و به یک داروی استاندارد تبدیل گردید، پژوهشگران از ارائه آن به شرکت‌کنندگان خودداری کردند. آن‌ها نه تنها درمان را دریغ کردند، بلکه فعالانه از دریافت درمان توسط شرکت‌کنندگان جلوگیری می‌نمودند. پژوهشگران با همکاری سازمان‌های محلی، اطمینان حاصل می‌کردند که هیچ یک از مردان تحت مطالعه تحت درمان با پنی‌سیلین قرار نگیرند [۵۸]. این وضعیت تا سال ۱۹۷۲ ادامه یافت، زمانی که خبرگزاری آسوشیتدپرس، گزارشی از این مطالعه منتشر کرد و افکار عمومی را با واقعیت چهل سال نظارت بر مرگ تدریجی صدها مرد سیاه‌پوست روبه‌رو ساخت [۵۵]. در طول این مدت، ده‌ها نفر از شرکت‌کنندگان بر اثر عوارض سفلیس جان‌باخته بودند، همسران و فرزندان آن‌ها نیز به این بیماری مبتلا شده بودند، و پژوهشگران همچنان به گردآوری داده‌ها ادامه می‌دادند. از منظر تاریخ اخلاق آماری، مطالعه توسسگی پرسش‌هایی اساسی را پیش می‌کشد. پژوهشگران این مطالعه از روش‌های آماری دقیق و پیشرفته‌ای استفاده می‌کردند. آن‌ها داده‌ها را با دقت ثبت می‌کردند، تحلیل‌های آماری معتبری انجام می‌دادند، و یافته‌های خود را در مجلات معتبر پزشکی منتشر می‌نمودند. در طول چهل سال، بیش از ده‌ها مقاله از این مطالعه در مجلات به چاپ رسید [۳۱]. نکته تأمل‌برانگیز این است که هیچ یک از منتقدان هم‌عصر این مطالعه، به روش‌شناسی آماری آن ایراد نگرفتند. ایراد اصلی، نه در روش‌های تحلیل داده، بلکه در طراحی پژوهش و نبود رضایت آگاهانه بود. اما این دو حوزه، یعنی طراحی و اخلاق گردآوری داده، تا پیش از رسوایی توسسگی به‌ندرت در حیطه مسئولیت آماردانان تعریف می‌شدند [۴۷]. آماردانان مشارکت‌کننده در این پروژه، خود را تحلیلگران فنی می‌دانستند که وظیفه‌شان تنها پردازش داده‌های گردآوری

1. Macon County
2. Alabama

3. Belmont Report

مطالعه توسکگی دست کم سه درس اساسی برای تاریخ اخلاق آماری به همراه داشت [۴۷، ۲۴]:

نخست: روش شناسی صحیح، فقدان اخلاق در طراحی پژوهش را جبران نمی کند. آماردانی که داده های حاصل از مطالعه ای غیر اخلاقی را تحلیل می کند، نمی تواند باتکیه بر دقت فنی تحلیل خود، مسئولیت اخلاقی را نفی کند.

دوم: مسئولیت اخلاقی آماردان از مرحله تحلیل فراتر می رود و به مرحله طراحی پژوهش و تعریف مسئله نیز تسری می یابد.

سوم: اخلاق پژوهش نیازمند نهادهای بازدارنده است، نه صرفاً وجدان فردی. مطالعه توسکگی چهل سال ادامه یافت نه به دلیل فقدان آگاهی اخلاقی در پژوهشگران منفرد، بلکه به دلیل فقدان سازوکارهای نهادی برای نظارت، بازخواست و توقف پژوهش های غیر اخلاقی.

تاریخ اصلاح نژادی و جنایات علمی قرن بیستم نشان می دهد که چگونه نظریه ای که ابتدا در محافل علمی محترم تلقی می شد، به ابزاری برای سرکوب و نسل کشی تبدیل گردید. این درس ها، مسیر تاریخ اخلاق آماری را در دهه های بعدی تعیین کردند و جامعه جهانی را به بازاندیشی در نسبت علم با مسئولیت اجتماعی فراخواندند.

۴. نهادهای سازی اخلاق در آمار

گزارش بلومونت و تأسیس هیئت های بازبینی نهادی تحول مهمی در تاریخ اخلاق پژوهش بود، اما تأثیر آن بر جامعه آماردانان تدریجی و همراه با چالش های خاص خود بود. دهه های ۱۹۸۰ تا ۲۰۰۵ دوره گذار از اخلاق فردی و وجدانی به اخلاق نهادی و حرفه ای در آمار محسوب می شود. در این دوره، انجمن های حرفه ای آمار نخستین استانداردهای اخلاق آماری را تدوین کردند، آماردانان در مناقشات اخلاقی بخش دولتی و صنعتی نقش فعال تری یافتند، و دامنه مسئولیت اخلاقی از صداقت فنی به مسئولیت در برابر جامعه گسترش یافت [۴۷].

۴-۱. پیشینه تدوین راهنماهای اخلاق آماری

پیش از آنکه انجمن آمار آمریکا در سال ۱۹۸۱ نخستین راهنمای اخلاقی خود را منتشر کند، نزدیک به سه دهه تلاش مستمر اما نافرجام برای تدوین یک شیوه نامه رفتار حرفه ای صورت گرفته بود. بررسی این پیشینه تاریخی که در مرجع [۴۴] به خوبی مستند شده، علل تأخیر در نهادهای سازی اخلاق آماری را روشن می سازد و نشان می دهد تدوین استانداردهای اخلاقی همواره با چالش های

نظری و نهادی عمیقی روبه رو بوده است. نخستین گام رسمی در این مسیر به سال ۱۹۴۹ بازمی گردد، زمانی که انجمن آمار آمریکا پیشنهاد کرد که تدوین یک شیوه نامه رفتار اخلاقی بخشی از وظایف کمیسیون استانداردها و سازماندهی آماری این انجمن باشد که در همان زمان تشکیل شده بود. سه سال بعد، در سال ۱۹۵۲، هیئت مدیره انجمن به رئیس وقت، ویلیام کوکران^۱، اختیار داد تا کمیته ای موقت برای انجام این مهم تشکیل دهد. ریاست این کمیته به رنسیس لیکرت^۲ سپرده شد [۴۴]. کمیته لیکرت در سال ۱۹۵۴ گزارشی منتشر کرد که در آن بر سه اصل تأکید داشت: نخست، انجمن باید به سمت تدوین بیانیه ای از استانداردهای مورد توافق (هم فنی و هم اخلاقی) حرکت کند؛ دوم، کمیته باید به اصول کلی بپردازد نه جزئیات موردی؛ و سوم، فرایند تدوین باید به گونه ای طراحی شود که بیانیه نهایی از حمایت اکثریت قاطع اعضای انجمن برخوردار باشد. کمیته به طور مشخص توصیه کرد که پیش از هر اقدامی، نظرسنجی از اعضا برای سنجش میزان تمایل به تدوین این استانداردها انجام شود. همچنین پیشنهاد شد که انجمن از رویه ای مشابه انجمن روان شناسی آمریکا پیروی کند (۱) استفاده از روش رویدادهای بحرانی برای شناسایی مسائل که اعضا در عمل با آن مواجه شده اند و سنجش میزان اجماع پیرامون آنها؛ و (۲) اتکا بر مشارکت گسترده اعضا در فرایند تدوین. تأکید کمیته بر این بود که استانداردها باید از تجربیات اعضا استخراج شوند، نه آنکه کمیته از تجربه محدود خود بیانیه ای را تهیه کند [۴۴]. هم زمان با فعالیت این کمیته، بحث های نظری مهمی درباره ماهیت و محتوای استانداردهای اخلاق آماری در نشست های انجمن و صفحات نشریات آماری در گرفت. به عنوان مثال، کورت^۳ در مقاله خود [۳۵] تأکید می کند: از آنجا که استانداردهای رفتار حرفه ای به تمامی فعالیت های انسانی تعمیم دارند، باید بر جزئیات عملی متمرکز باشند، نه آنکه صرفاً به کلی گویی هایی چون صداقت و وفاداری بسنده کنند. مورتون^۴ با تفکیک آماردانان به سه گروه: نظریه پرداز (نیاز اندک به استاندارد اخلاقی)، متخصص موضوعی (نیاز متوسط به استاندارد اخلاقی)، و آماردان گردآورنده (نیاز زیاد به استاندارد اخلاقی)، نشان داد که شدت چالش های اخلاقی در هر گروه متفاوت است و بنابراین یک استاندارد اخلاقی واحد، ممکن است برای همه گروه ها مناسب نباشد [۶۶]. براون^۵ در مقاله ای تأثیر گذار [۳۲] سه گونه استاندارد اخلاقی را از هم متمایز کرد. گونه نخست، الگوبرداری از سوگند بقراط بود که به اعتقاد او برای آماردانان مناسب نیست، زیرا پزشک با مریضی مواجه است که خواهان بهبودی است، در حالی که آماردان اغلب با کارفرمایانی سروکار دارد که گاهی خواهان نتایجی خلاف واقعیت هستند. گونه دوم، استانداردهای خودپسندانه ای بودند که براون آن ها را نابخردانه

1. William Cochran
2. Rensis Likert
3. Andrew T. Court

4. Joseph E. Morton
5. Theodore H. Brown

۴-۲. نخستین راهنمای اخلاق آماری

پس از دوره آزمایشی، انجمن آمار آمریکا در سال ۱۹۸۳ نخستین راهنمای اخلاقی برای کارهای آماری را به طور رسمی منتشر کرد. این راهنما که از آن زمان تاکنون چندین بار بازنگری شده، شش حوزه اصلی مسئولیت اخلاقی آماردانان را مشخص می‌کرد [۲۱]:

۱. **مسئولیت در قبال آزمودنی‌ها و ذی‌نفعان:** مسئولیت در قبال آزمودنی‌ها و ذی‌نفعان: حفاظت از حریم خصوصی، رعایت محرمانگی داده‌ها، و اطمینان از اخذ رضایت آگاهانه.

۲. **مسئولیت در قبال کارفرما و کارگزار:** ارائه تحلیل‌های صادقانه و پرهیز از تطبیق نتایج با خواست کارفرما.

۳. **مسئولیت در قبال جامعه:** توجه به پیامدهای اجتماعی کاربرد یافته‌های آماری و پرهیز از مشارکت در پروژه‌های غیراخلاقی.

۴. **مسئولیت در قبال روش‌شناسی:** به‌کارگیری روش‌های مناسب، گزارش محدودیت‌ها، و پرهیز از سوگیری‌های ناخواسته.

۵. **مسئولیت در قبال همکاران و حرفه:** نقد منصفانه، تصدیق مشارکت‌ها، و پرهیز از تضاد منافع.

۶. **مسئولیت در قبال آموزش و نسل آینده:** تربیت آماردانانی که ارزش‌های اخلاقی را درونی کرده‌اند.

انجمن بین‌المللی آمار نیز در ۱۹۸۵ راهنمای اخلاق حرفه‌ای خود را منتشر کرد. این اعلامیه که رویکردی جهانی‌تر داشت، بر تفاوت‌های فرهنگی در تعریف اخلاق تأکید می‌کرد و درعین‌حال اصول اساسی چون صداقت، شفافیت، و مسئولیت اجتماعی را مشترک میان همه فرهنگ‌ها می‌دانست [۵۷]. در سال‌های بعد، هر دو نهاد راهنمای اخلاق آماری خود را بازنگری کردند. انجمن آمار آمریکا در سال ۲۰۱۸ نسخه بازنگری شده راهنمای اخلاقی را منتشر کرد که توجه بیشتری به چالش‌های عصر داده‌های بزرگ و یادگیری ماشین داشت [۲۳].

۴-۳. آمار در دولت و صنعت

در همین دوره، آماردانان شاغل در بخش‌های دولتی و صنعتی با چالش‌های اخلاقی متفاوتی نسبت به همکاران دانشگاهی خود مواجه بودند. اگر در محیط دانشگاهی مسئله اصلی اخذ رضایت آگاهانه و حفاظت از آزمودنی‌ها بود، در بخش دولتی و صنعتی چالش‌های دیگری برجسته می‌شدند [۲۴]. آماردانان شاغل در سازمان‌های آماری دولتی (مانند اداره آمار آمریکا یا دفتر آمار کانادا) گاه تحت فشار قرار می‌گرفتند تا روش‌های تحلیلی یا

می‌خواند، زیرا یا فراموش خواهند شد، یا نادیده گرفته خواهند شد، یا عادات آماردانان را به انعطاف‌ناپذیری محدود خواهند کرد. گونه سوم، الگوی کتاب مقدس بود: چند آرمان کلی همراه با مثال‌های فراوان، نه مجموعه‌ای از قواعد جزئی برای تمامی شرایط. براون شش آرمان را پیشنهاد کرد که از آن جمله می‌توان به اشتیاق برای یافتن حقیقت، تلاش مداوم برای ارائه روشن دستاوردهای خود، و جستجوی خستگی‌ناپذیر برای یافتن شواهد در تأیید یک فرضیه، نه در تأیید یک نتیجه از پیش تعیین شده اشاره کرد. فریمن^۱ نیز در همین دوره پیشنهاد سوگند آماردان را مطرح کرد که در آن آماردان خود را جستجوگر حقیقت معرفی می‌کند و متعهد می‌شود که در میان هیاهوی بازار یا میدان سیاست، فردی نباشد که ارقام را برای اثبات نتیجه‌ای ناهمخوان برگزیند. او همچنین یازده ماده مشخص را در حمایت از این سوگند پیشنهاد کرد که از آن جمله می‌توان به پرهیز از طراحی آزمایش‌های سوگیرانه، حفظ ناشناس بودن پاسخ‌دهندگان، و جلوگیری از سوءاستفاده از ارقام تهیه شده اشاره کرد [۵۰]. پیرو توصیه کمیته لیکرت، دو نظرسنجی برای سنجش تمایل اعضا به تدوین استانداردهای اخلاق آماری انجام شد. نظرسنجی مقدماتی که برای شعبه بوستون انجمن آمار انجام شد، نشان داد که تنها حدود یک‌سوم از اعضای آن به شدت طرفدار تدوین این استانداردها هستند. نظرسنجی دوم که تمامی شعبه فعال انجمن در اوایل دهه ۱۹۵۰ را پوشش می‌داد، نشان داد که تنها ۲۱ شعبه از تدوین استانداردها حمایت می‌کنند. کمیته نظرسنجی درباره استانداردها، در سال ۱۹۵۶ گزارش خود را منتشر کرد و بر اساس یافته‌های نظرسنجی، به انجمن توصیه نمود که تا زمانی که علاقه بیشتری به تدوین استانداردها وجود ندارد، اقدام را به تعویق بیندازد [۴۴]. با این توصیه، تلاش‌های رسمی انجمن برای تدوین یک مجموعه استاندارد حرفه‌ای برای دو دهه متوقف شد. علاقه به تدوین استانداردهای اخلاق آماری در اواخر دهه ۱۹۷۰ بار دیگر احیا شد. در سال ۱۹۷۶، جرزی نیمن^۲ طی نامه‌ای به هیئت‌مدیره، موضوع اخراج اعضای که عمداً داده جعل می‌کنند را مطرح کرد. در همان سال، سی. ترنس ایرلند^۳ نیز طی نامه‌ای به مدیر اجرایی و اعضای هیئت‌مدیره، خواستار تشکیل کمیته‌ای برای بررسی مسائل اخلاقی شد. در ژانویه ۱۹۷۷، کمیته موقتی برای تدوین راهنمای اخلاقی انجام کارهای آماری، به ریاست ترنس ایرلند تشکیل شد. این کمیته طی سه سال بعد، نسخه‌ای از این راهنما را تهیه و دو بار (در سال‌های ۱۹۸۰ و ۱۹۸۱) در خبرنامه انجمن منتشر کرد تا نظرات اعضا و شعب انجمن را دریافت کند. سرانجام در اوت ۱۹۸۱، هیئت‌مدیره انجمن نسخه نهایی «راهنمای اخلاقی برای کارهای آماری»^۴ را برای یک دوره آزمایشی سه‌ساله تصویب کرد [۴۴، ۲۱].

1. Freeman
2. Jerzy Neyman
3. C. Terence Ireland

4. Ethical Guidelines for Statistical Practice
5. International Statistical Institute (ISI)

یکی دیگر از حوزه‌هایی که در این دوره گسترش یافت، مسئولیت آماردان در قبال کاربردهای ثانویه داده‌ها بود. با گسترش پایگاه‌های داده بزرگ و افزایش دسترسی به داده‌های ثانویه، این پرسش مطرح شد که آیا آماردانی که داده‌های حساس (مانند داده‌های پزشکی یا مالی) را برای هدف خاصی جمع‌آوری کرده، مسئول کاربردهای بعدی آن داده‌ها توسط دیگران نیز هست؟ پاسخ به این پرسش همچنان محل بحث است، اما در استانداردهای اخلاق آماری این دوره، تأکید بر مدیریت مسئولانه داده‌ها به‌عنوان بخشی از وظایف اخلاقی آماردان گنجانده شد [۲۴].

۴-۵. کاستی‌های نهادهای اخلاقی

باوجود پیشرفت‌های چشمگیر در نهادینه‌سازی اخلاق آماری در این دوره، کاستی‌هایی نیز وجود داشت. مهم‌ترین محدودیت، ضعف سازوکارهای اجرایی بود. استانداردهای اخلاقی انجمن‌های حرفه‌ای عمدتاً توصیه‌ای بودند و سازوکار مؤثری برای تنبیه متخلفان یا بازخواست نهادهایی که فشارهای غیراخلاقی بر آماردانان وارد می‌کردند، فراهم نمی‌آوردند. آماردانی که به دلیل مقاومت در برابر فشار کارفرما شغل خود را از دست می‌داد، در عمل پناهگاه نهادی مؤثری نداشت [۴۷]. علاوه بر این، تمرکز اصلی استانداردهای اخلاقی این دوره بر رفتار فردی آماردان بود و کمتر به ساختارهای نهادی و فشارهای سیستمی که منشأ بسیاری از چالش‌های اخلاقی بودند، توجه می‌کرد. این کاستی در دهه بعد و با ظهور بحران تکرارپذیری آشکار شد. بحرانی که نشان داد حتی در حضور استانداردهای اخلاق آماری و هیئت‌های بازبینی نهادی، فشارهای سیستمی برای تولید نتایج معنادار می‌تواند به فساد تدریجی رویه‌های آماری در سطح وسیع منجر شود. بخش بعدی این مقاله به بررسی این بحران و پیامدهای آن برای اخلاق آماری اختصاص دارد.

۵. بحران تکرارپذیری

اگر مطالعه توسکگی نشان داد که بی‌توجهی به اخلاق در طراحی پژوهش می‌تواند به فاجعه‌ای انسانی منجر شود، بحران «تکرارپذیری»^۳ (توانایی تکرار نتایج یک پژوهش با استفاده از همان داده‌ها و روش‌های تحلیلی) که در دهه ۲۰۱۰ دامن‌گیر علوم تجربی به ویژه روان‌شناسی شد، افشاگر بحران‌های درونی خود علم آمار بود. این بار مسئله نه سوءرفتار آشکار یا پژوهش‌های غیرانسانی، بلکه فشارهای نهادی و رویه‌های تحلیلی نادرست بود که در سطح وسیع و در درون جریان اصلی علم، به تولید نتایجی غیرقابل اعتماد انجامید. بحران تکرارپذیری، آمار را از مقام ابزار بی‌طرف کشف حقیقت به عرصه‌ای از مناقشات اخلاقی-روش‌شناختی فرود آورد و نهادهای بین‌المللی آمار را واداشت تا رسماً به بازتعریف مسئولیت‌های اخلاقی آماردانان بپردازند [۲۲، ۷۱].

نحوه گزارش‌دهی را به‌گونه‌ای تغییر دهند که نتایج با اولویت‌های سیاسی دولت تطبیق کند. یکی از مشهورترین موارد این فشارها در اوایل دهه ۱۹۹۰ در دفتر آمار کانادا رخ داد، زمانی که دولت محافظه‌کار خواستار تغییر روش محاسبه نرخ بیکاری شد تا آمارهای منتشرشده مطلوب‌تر به نظر برسند. آماردانان ارشد این سازمان در اقدامی بی‌سابقه، به‌صورت علنی در برابر این فشار مقاومت کردند و استعفای جمعی را تهدید نمودند. این واقعه منجر به تدوین اصل استقلال حرفه‌ای در آمار دولتی شد که امروزه در بسیاری از کشورها به‌عنوان یک هنجار نهادی پذیرفته شده است [۴۶].

در بخش صنعتی، آماردانان با چالش تضاد منافع میان وظیفه حرفه‌ای برای گزارش صادقانه نتایج و فشار کارفرما برای ارائه تصویری خوش‌بینانه از محصول یا فرایند مواجه بودند. صنایع داروسازی نمونه بارز این چالش بود. آماردانان شاغل در شرکت‌های داروسازی در دهه ۱۹۸۰ و ۱۹۹۰ در موارد متعددی شاهد تلاش برای دست‌کاری تحلیل‌های آماری به نفع محصولات شرکت بودند. این تجربیات منجر به شکل‌گیری مفهوم استقلال آماری در قراردادهای کاری شد: شرطی که بر اساس آن آماردان حق دارد نتایج واقعی تحلیل را بدون سانسور، یا تغییر توسط کارفرما منتشر کند [۳۹].

۴-۴. گسترش مسئولیت اخلاقی

در این دوره، درک آماردانان از مسئولیت اخلاقی خود دستخوش تحول اساسی شد. در گذشته، تعریف غالب از اخلاق در آمار محدود به صداقت فنی بود: پرهیز از جعل داده، انتخاب آگاهانه روش‌های تحلیلی، و گزارش صادقانه نتایج. اما تجربیات دهه‌های ۱۹۸۰ و ۱۹۹۰ نشان داد که صداقت فنی به‌تنهایی کافی نیست [۴۷]. در اواخر دهه ۱۹۸۰، پژوهشگران حوزه فراتحلیل به‌تدریج متوجه پدیده‌ای شدند که بعدها «سوگیری انتشار»^۱ نام گرفت: مطالعاتی که نتایج «معنادار آماری»^۲ دارند (یعنی نتیجه‌ای که از پژوهش به دست آمده، به احتمال زیاد شانس نیست) در مقایسه با پژوهش‌هایی که نتایج غیرمعنادار دارند، شانس بیشتری برای انتشار در مجلات علمی پیدا می‌کنند [۷۶]. این سوگیری که ریشه در ساختارهای نهادی علم داشت (فشار برای انتشار، داوری هم‌تا که به نتایج معنادار اهمیت می‌دهد، و غیره)، آماردانان را با پرسش اخلاقی تازه‌ای روبه‌رو کرد: آیا آماردانی که نتایج غیرمعنادار پژوهش خود را بایگانی می‌کند و برای انتشار آن تلاش نمی‌کند، مرتکب کوتاهی اخلاقی شده است؟ پاسخ تدریجی که در استانداردهای اخلاقی بعدی منعکس شد، مثبت بود: آماردانان وظیفه دارند، مستقل از معناداری یا عدم معناداری نتایج، برای انتشار یافته‌های خود تلاش کنند تا تصویر کامل و بدون سوگیری از واقعیت در دسترس جامعه علمی قرار گیرد.

1. Publication bias
2. Statistically significant

3. Reproducibility

۵-۱. نقش آمار در بحران

نشانه‌هایی از بحران تکرارپذیری از دهه‌ها پیش وجود داشت. در اوایل دهه ۱۹۹۰، جان یونیدیس^۱، اپیدمیولوژیست یونانی-آمریکایی، در مقاله‌ای تأثیرگذار با عنوان «چرا بیشتر نتایج پژوهش‌های منتشرشده غلط از آب درمی‌آیند؟» نشان داد که به دلیل ترکیبی از عوامل از جمله حجم نمونه‌های کوچک، اندازه اثر^۲ (میزان تأثیر یک متغیر بر متغیر دیگر، مستقل از حجم نمونه) و سوگیری انتشار، بخش بزرگی از یافته‌های علوم زیستی و رفتاری ممکن است غیرقابل اعتماد باشند [۵۶]. پیش از آن، در دهه ۱۹۹۰، روان‌شناسان تجربی هشدار داده بودند که یافته‌های برجسته حوزه روان‌شناسی اجتماعی غالباً در تلاش‌های تکرارپذیری با شکست مواجه می‌شوند [۷۶]. اما این هشدارها تا اوایل دهه ۲۰۱۰ جدی گرفته نشدند. نقطه عطف، انتشار نتایج پروژه تکرارپذیری روان‌شناسی در سال ۲۰۱۵ بود. این پروژه که با مشارکت دهه‌آزمایشگاه در سراسر جهان انجام شد، تلاش کرد ۱۰۰ مطالعه برجسته منتشرشده در سه مجله معتبر روان‌شناسی را تکرار کند. نتیجه شوک‌آور این بود که فقط ۳۹ درصد از تکرارها نتایج معناداری مشابه با مطالعات اصلی نشان دادند. اندازه اثر در تکرارها به طور متوسط نصف اندازه اثر گزارش شده در مطالعات اصلی بود [۷۱]. این یافته نه فقط روان‌شناسی، بلکه تمامی علوم تجربی و به‌ویژه جامعه‌آماردانان را تحت تأثیر قرار دارد. اگر یافته‌های منتشرشده در مجلات معتبر تا این حد غیرقابل تکرارند، آیا مشکل صرفاً در روش‌های روان‌شناسی است، یا ریشه در شیوه‌های رایج تحلیل آماری دارد؟ پاسخ به این پرسش، انگشت اتهام را به‌سوی سوءاستفاده از روش‌های آماری به‌ویژه آزمون فرضیه‌های معناداری نشانه گرفت. در تحلیل‌های آماری، کمیت p -مقدار^۳، احتمال به‌دست آمدن نتایج مشاهده‌شده یا نتایجی افراطی‌تر از آن، تحت درست بودن فرضیه صفر^۴ (که بر اساس آن هیچ تفاوت یا رابطه‌ای بین متغیرها وجود ندارد) را نشان می‌دهد. به عبارت ساده‌تر، این کمیت می‌گوید: اگر در واقعیت هیچ تفاوت یا رابطه‌ای وجود نداشته باشد، چقدر احتمال دارد که داده‌ها به طور تصادفی چنین تفاوتی را نشان دهند؟ هرچه p -مقدار کوچک‌تر باشد، شواهد قوی‌تری بر ضد فرضیه صفر وجود دارد. معمولاً به طور قراردادی آستانه ۰/۰۵ به‌عنوان مرز معناداری آماری در نظر گرفته می‌شود. پژوهشگران حوزه فراتحلیل نشان دادند که فشار نهادی برای تولید نتایج معنادار آماری، منجر به شیوع رفتارهایی شده که بعدها «هک p -مقدار»^۵ نام گرفت. این اصطلاح به مجموعه‌ای از شیوه‌های تحلیلی فرصت‌طلبانه اشاره دارد که اگرچه هریک به‌تنهایی ممکن است از نظر فنی مجاز باشند، اما در مجموع احتمال دستیابی به نتایج معنادار آماری را به شدت افزایش می‌دهند. برخی از این شیوه‌ها عبارت‌اند از [۷۷]:

توقف اختیاری: ادامه گردآوری داده تا رسیدن به معناداری آماری انتخاب متغیرهای وابسته پس از مشاهده داده‌ها: آزمون چندین متغیر وابسته و گزارش تنها آن‌هایی که نتایج معنادار دارند.

حذف نقاط دورافتاده پس از تحلیل: حذف مشاهدات نامطلوب به بهانه دورافتاده بودن پس از مشاهده تأثیر آن‌ها بر نتایج.

افزودن متغیرهای کنترل به شیوه فرصت‌طلبانه: افزودن متغیرهای کنترل به مدل تا زمانی که نتایج معنادار شوند.

تغییر در تعریف متغیرهای اصلی: تغییر متغیر وابسته اصلی پس از مشاهده نتایج.

این شیوه‌ها که در ادبیات تخصصی با عنوان انعطاف‌پذیری تحلیلی شناخته می‌شوند [۵۲]، به پژوهشگر این امکان را می‌دهند که از میان چندین روش ممکن برای تحلیل داده‌ها، راهی را انتخاب کند که به نتیجه معنادار منجر می‌شود. نتیجه نهایی، افزایش کاذب نرخ خطای نوع اول (رد اشتباه فرضیه صفر یا اعلام وجود تفاوت یا رابطه درحالی‌که در واقعیت وجود ندارد) و تولید یافته‌هایی است که در پژوهش‌های بعدی قابل تکرار و تأیید نیستند.

۵-۲. واکنش نهادهای به بحران تکرارپذیری

نکته حائز اهمیت از منظر اخلاق آماری این است که بسیاری از پژوهشگران مرتکب هک p -مقدار با آگاهی و عمد نمی‌شدند. آن‌ها صرفاً در چارچوب فشارهای نهادی عمل می‌کردند: برای استخدام و ارتقای مرتبه به مقاله نیاز دارند، مجلات معتبر تمایل به انتشار نتایج معنادار دارند، و راهنماهای روش‌شناختی نیز به‌روشنی مشخص نکرده بودند که کدام یک از این شیوه‌ها غیراخلاقی هستند. به عبارت دیگر، ساختارهای نهادی علم به‌گونه‌ای طراحی شده بودند که رفتارهای فرصت‌طلبانه را تشویق می‌کردند [۶۸]. این وضعیت، اخلاق آماری را از حیطه رفتار فردی به حیطه بازطراحی نهادی وارد کرد. پرسش‌های تازه‌ای مطرح شدند:

• آیا دآوری هم‌تا باید نتایج غیرمعنادار را نیز به یک اندازه شایسته انتشار بداند؟

• آیا مجلات علمی باید طرح‌های پژوهشی را ملزم کنند تا میان «تحلیل تأییدی»^۶ (تحلیلی که برای آزمون فرضیه‌های از پیش تعیین‌شده انجام می‌شود) و «تحلیل اکتشافی»^۷ (تحلیلی که برای کشف الگوها و فرضیه‌های جدید در داده‌ها انجام می‌شود) تمایز قائل شوند؟

• آیا آماردانان به‌عنوان دروازه‌بانان روش‌شناختی، مسئولیتی در قبال آموزش و نظارت بر این شیوه‌ها دارند؟

1. John Ioannidis
2. Effect size
3. p-value
4. Null hypothesis

5. p-hacking
6. Confirmatory Analysis
7. Exploratory Analysis

بحران تکرارپذیری سرانجام برخی از نهادهای آماری را به واکنشی بی‌سابقه واداشت. در مارس ۲۰۱۶، انجمن آمار آمریکا بیانیه‌ای دربارهٔ p -مقدار منتشر کرد. این بیانیه که با همکاری گروهی از آماردانان برجسته تدوین شد، شش اصل را اعلام کرد که نقشی تعیین‌کننده در بازتعریف جایگاه اخلاقی آماردانان محسوب می‌شود [۲۲]:

۱. p -مقدار نمی‌تواند نشان دهد که فرضیهٔ صفر درست است یا اینکه یک اثر از نظر علمی مهم است.

۲. p -مقدار به‌تنهایی نمی‌تواند اندازهٔ یک اثر یا اهمیت یک نتیجه را نشان دهد.

۳. وجود یک تفاوت معنادار آماری لزوماً به معنای وجود یک تفاوت علمی یا عملی مهم نیست.

۴. انتخاب آستانهٔ 0.05 برای معناداری، یک قرارداد است و مبنای علمی یا فلسفی ضروری ندارد.

۵. p -مقدار به‌تنهایی نمی‌تواند معیار کافی برای ارزیابی یک فرضیه یا مدل باشد.

۶. p -مقدار به‌تنهایی نمی‌تواند شواهد موجود در داده‌ها را به طور کامل خلاصه کند.

اهمیت تاریخی این بیانیه در دو نکته بود. نخست، انجمن آمار آمریکا به‌عنوان یکی از مهم‌ترین نهادهای حرفه‌ای بین‌المللی در حوزه آمار، رسماً اذعان کرد که سوءاستفاده از p -مقدار صرفاً یک مسئله فنی نیست، بلکه یک مسئله اخلاقی است. دوم، این بیانیه نشان داد که جامعهٔ آماردانان دیگر نمی‌تواند خود را ناظری بی‌طرف و منفعل در برابر شیوه‌های تحلیلی نادرست بداند. آماردانان وظیفه دارند در فرایندهای طراحی پژوهش، تحلیل داده و داوری مقالات، به طور فعال از تکرار رویه‌های فرصت‌طلبانه جلوگیری کنند. بحران تکرارپذیری و به دنبال آن بیانیهٔ انجمن آمار آمریکا، پیامدهای مهمی برای اخلاق آماری به همراه داشت که تا امروز ادامه دارد [۶۸، ۶۷]. یکی از مهم‌ترین این پیامدها، گسترش رویکرد «پیش‌ثبت نویسی» و «پروپوزال نویسی» بود. در این رویکرد، پژوهشگر موظف است پیش از دسترسی و مشاهده داده‌ها، فرضیه‌ها، متغیرها و روش تحلیل را ثبت کند تا تمایز میان تحلیل‌های تأییدی و تحلیل‌های اکتشافی حفظ شود [۶۷]. بسیاری از مجلات معتبر در علوم پزشکی و روان‌شناسی، پیش‌ثبت نویسی را به یکی از شروط انتشار مطالعات تجربی تبدیل کردند. این رویکرد، هم به‌عنوان سندی برای شفافیت پژوهش عمل می‌کند و هم به پژوهشگر کمک می‌کند تا مسیر پژوهش را تعیین کند. همچنین به خواننده (داور، استاد راهنما، یا مخاطب مقاله) نشان می‌دهد که کدام بخش‌های پژوهش از پیش برنامه‌ریزی شده و کدام بخش‌ها در حین پژوهش کشف شده‌اند.

تحول مهم دیگر، فاصله‌گرفتن تدریجی از اتکای صرف به آزمون معناداری و حرکت به‌سوی گزارش اندازه اثر، فاصله اطمینان و تحلیل‌های قوی‌تر بود [۳۶]. این تغییر، بازتاب نوعی دگرگونی معرفت‌شناختی در فهم شواهد آماری بود؛ دگرگونی که بر عدم قطعیت، اندازه اثر و بازتولیدپذیری^۱ (توانایی تولید نتایج مشابه با استفاده از داده‌های جدید و روش‌های مشابه) تأکید بیشتری داشت. بحران تکرارپذیری همچنین بر آموزش آمار اثر گذاشت. روشن شد که آموزش صرفاً فنی روش‌های آماری، بدون توجه به زمینه‌های تاریخی، فلسفی و اخلاقی آن‌ها، ناکافی است. از این‌رو، در بسیاری از دانشگاه‌ها مباحثی مانند اخلاق پژوهش کمی، بازتولیدپذیری و فلسفه استنباط آماری وارد برنامه‌های آموزشی تحصیلات تکمیلی شد [۵۳]. در سال‌های بعد، مفهوم تکرارپذیری نیز خود به موضوعی نظری تبدیل شد. برخی پژوهشگران میان تکرارپذیری فنی (دستیابی به نتایج یکسان با همان داده‌ها)، تکرارپذیری تحلیلی (کسب نتایج مشابه با تحلیل مستقل داده‌های مشابه)، و تکرارپذیری علمی (تولید یافته‌های مشابه در پژوهش‌های مستقل) تمایز قائل شدند [۷۳]. هر یک از این سطوح، مسائل اخلاقی و روش‌شناختی متفاوتی را مطرح می‌کنند. با این حال، بحران تکرارپذیری پایان چالش‌های اخلاق آماری نبود. هم‌زمان با گسترش داده‌های بزرگ (مه داده)^۲، یادگیری ماشین و هوش مصنوعی، مسائل تازه‌ای درباره حریم خصوصی، سوگیری الگوریتمی، شفافیت مدل‌ها و عدالت آماری پدید آمد؛ مسائلی که بسیاری از مفروضات سنتی اخلاق پژوهش را به چالش کشیده‌اند. از این‌رو، اخلاق آماری در قرن بیست و یکم بیش از هر زمان دیگر به مسئله‌ای میان‌رشته‌ای و مرتبط با سیاست‌گذاری علم و فناوری تبدیل شده است.

۶. عصر داده‌های بزرگ

هم‌زمان با فروکش کردن نسبی غبار بحران تکرارپذیری، تحول دیگری در حال وقوع بود که مرزهای اخلاق آماری را فراتر از چارچوب‌های سنتی گسترش داد. عصر داده‌های بزرگ و یادگیری ماشین پرسش‌های تازه‌ای را پیش کشید که پاسخ به آن‌ها با اصول منشور اخلاقی نورنبرگ و گزارش بلومونت (که عمدتاً برای کارآزمایی‌های بالینی با آزمودنی‌های انسانی طراحی شده بودند) به‌سادگی ممکن نبود. این دوره، اخلاق آماری را با سه چالش اساسی روبه‌رو ساخت: بحران اخذ رضایت آگاهانه در مقیاس کلان، مسئلهٔ سوگیری الگوریتمی و عدالت آماری، و بازتعریف مسئولیت آماردان در زنجیرهٔ ارزش داده [۷۰، ۷۶].

۶-۱. بحران اخذ رضایت آگاهانه

مدل سنتی «اخذ رضایت آگاهانه»^۳ که ریشه در منشور اخلاقی نورنبرگ دارد، بر سه پایه استوار بود: آگاهی (آزمودنی از ماهیت،

1. Replicability
2. Big data

3. Informed consent

کنند. این یافته‌ها، اخلاق آماری را از حیطه رفتار حرفه‌ای فردی به حیطه عدالت اجتماعی وارد کرده‌اند [۷۰].

نمونه ۱: در ایالات متحده، سامانه‌هایی مانند COMPAS^۱ برای پیش‌بینی احتمال ارتکاب جرم در زندانیان و کمک به قضات در تصمیم‌گیری درباره وثیقه و آزادی مشروط، از الگوریتم‌های پیش‌بینی جرم استفاده می‌شود. تحقیقات نشان داد که این الگوریتم‌ها به طور نظام‌مند نرخ خطای مثبت بالاتری برای زندانیان سیاه‌پوست و نرخ خطای منفی بالاتری برای زندانیان سفیدپوست تولید می‌کنند. به تعبیر دیگر، خروجی الگوریتم حاکی از آن بود که زندانیان سیاه‌پوست با احتمالی بالاتر از مقدار واقعی مرتکب جرم مجدد می‌شوند، درحالی‌که این احتمال برای زندانیان سفیدپوست به میزان قابل‌توجهی پایین‌تر از واقعیت برآورد می‌گردد. این سوگیری، صرف‌نظر از دقت کلی الگوریتم (که معیاری آماری است)، نقض آشکار اصل عدالت در گزارش بلمونت بود [۱۹].

نمونه ۲: شرکت‌های بزرگ فناوری از الگوریتم‌های یادگیری ماشین برای غربالگری پرونده‌های متقاضیان کار استفاده می‌کنند. در سال ۲۰۱۸، گزارش شد که الگوریتم استخدام شرکت آمازون که بر اساس داده‌های تاریخی استخدام (که عمدتاً مرد بودند) آموزش‌دیده بود. الگوریتم، حضور نشانه‌های زنانه (از جمله نام یا واژگان مرتبط با زنان) در پرونده را به‌عنوان شاخصی برای کاهش شایستگی تعبیر می‌کرد و به طور نظام‌مند به متقاضیان زن امتیاز منفی می‌داد. [۳۷]. از دیدگاه آماری، این سوگیری‌ها ریشه در چند عامل دارند [۲۶]:

سوگیری داده‌های تاریخی: الگوریتم‌ها از داده‌های گذشته یاد می‌گیرند. اگر داده‌های تاریخی خود حامل سوگیریهای اجتماعی باشند، الگوریتم نیز همان سوگیری‌ها را به‌عنوان الگو، شناسایی و بازتولید می‌کند. نمونه بارز این مسئله، الگوریتم استخدامی است که بر اساس داده‌های مردسالارانه گذشته آموزش دیده است؛ چنین الگوریتمی، نه به دلیل ارزیابی واقعی توانمندی‌ها، بلکه به دلیل تحمیل الگوهای تاریخی، به‌طور نظام‌مند رزومه‌های زنان را رد می‌کند.

سوگیری در تعریف متغیر هدف: انتخاب متغیر هدف هرگز یک تصمیم خنثی و صرفاً فنی نیست؛ بلکه سرشار از ارزش‌گذاری است. اگر متغیرهایی چون خطر ارتکاب جرم یا شایستگی شغلی با سوگیری‌های نژادی، جنسیتی یا طبقاتی تعریف شوند، الگوریتم نیز همان سوگیری‌ها را به‌عنوان واقعیت بازتاب خواهد داد.

سوگیری در داده‌های برچسب‌گذاری: در یادگیری با نظارت^۵، برچسب‌هایی که به داده‌ها نسبت داده می‌شوند ممکن است خود

اهداف، خطرات و منافع پژوهش مطلع شود)، اختیار (رضایت بدون اجبار و باقابلیت انصراف در هر زمان داده شود)، و فردیت (رضایت مربوط به یک پژوهش مشخص با یک آزمودنی مشخص است). در عصر داده‌های بزرگ، هر سه پایه به چالش کشیده می‌شوند [۶۵]

چالش آگاهی: در پژوهش‌های مبتنی بر داده‌های بزرگ، اغلب داده‌ها از میلیون‌ها کاربر بدون آگاهی مستقیم آن‌ها گردآوری می‌شوند. کاربران اپلیکیشن‌ها، شبکه‌های اجتماعی، فروشگاه‌های آنلاین و حتی دستگاه‌های پوشیدنی مانند ساعت‌های هوشمند، روزانه حجم زیادی از داده‌های رفتاری، مکانی، زیستی و مالی تولید می‌کنند. این داده‌ها بعداً برای مقاصد پژوهشی (و نیز تجاری) مورد استفاده قرار می‌گیرند. اما آیا کاربری که هنگام نصب یک اپلیکیشن، شرایط و ضوابط پنجاه صفحه‌ای را بدون خواندن تأیید می‌کند، رضایت آگاهانه داده است؟ اکثر پژوهشگران حوزه اخلاق داده بر این باورند که مدل سنتی اخذ رضایت در مقیاس کلان عملاً غیرقابل اجرا است [۶۵].

چالش اختیار: در بسیاری از موارد، رضایت کاربران به داده‌هایشان نه یک انتخاب آزادانه، بلکه شرطی برای دسترسی به خدمات ضروری است. اگر برای استفاده از یک سرویس بانکی یا درمانی ناچار به پذیرش شرایط استفاده‌ای باشیم که اجازه استفاده ثانویه از داده‌ها را می‌دهد، آیا واقعاً اختیار داریم؟ این پدیده که «اجبار ساختاری»^۱ نام گرفته، یکی از چالش‌های اصلی اخلاق داده در عصر حاضر است [۲۶].

چالش فردیت: پژوهش‌های مبتنی بر داده‌های بزرگ اغلب از داده‌های ثانویه (داده‌هایی که برای هدف دیگری گردآوری شده‌اند) استفاده می‌کنند. در چنین پژوهش‌هایی، به دلیل انبوهی داده‌ها و عدم امکان شناسایی و دسترسی به تمامی افراد، رضایت‌گیری انفرادی عملاً غیرممکن است. آیا استفاده ثانویه از این داده‌ها بدون رضایت مجاز است؟ پاسخ به این پرسش در حوزه‌های مختلف متفاوت است. در اتحادیه اروپا، مقررات عمومی حفاظت از داده^۲ که در ۲۰۱۸ به اجرا درآمد، رویکردی سخت‌گیرانه اتخاذ کرده و استفاده ثانویه از داده‌های شخصی را بدون رضایت صریح ممنوع می‌کند، مگر در موارد استثنایی با منفعت عمومی روشن [۴۵].

۶-۲. سوگیری الگوریتمی

اگر بحران رضایت آگاهانه به ورودی‌های سامانه‌های داده‌محور مربوط می‌شود، مسئله «سوگیری الگوریتمی»^۳ به خروجی‌های آن‌ها مرتبط است. در دهه گذشته، انبوهی از مطالعات نشان داده‌اند که الگوریتم‌های یادگیری ماشین (که بر پایه روش‌های آماری ساخته شده‌اند) می‌توانند نابرابری‌های اجتماعی را بازتولید یا حتی تشدید

1. Structural coercion

2. General Data Protection Regulation (GDPR)

3. Algorithmic bias

4. Correctional Offender Management Profiling for Alternative Sanctions (COMPAS)

5. Supervised learning

تمامی مراحل زنجیره ارزش داده حضور فعال داشته باشد [۴۷،۲۳]

طراحی جمع‌آوری داده: آماردان باید در مرحله طراحی جمع‌آوری داده، اطمینان حاصل کند که فرایند اخذ رضایت آگاهانه (با در نظر گرفتن محدودیت‌های عصر داده‌های بزرگ) به درستی طراحی شده و گروه‌های آسیب‌پذیر مورد حمایت قرار گرفته‌اند.

مدیریت و حاکمیت داده: آماردان مسئول است که داده‌ها به گونه‌ای مدیریت شوند که حریم خصوصی افراد حفظ شود. روش‌هایی مانند ناشناس‌سازی و رمزنگاری، برای حفظ حریم خصوصی افراد، بخشی از ابزارهای اخلاقی آماردان امروزی هستند [۴۲].

طراحی الگوریتم و انتخاب معیار: آماردان باید در فرایند طراحی الگوریتم، ضمن توجه به دقت آماری، پیامدهای توزیعی الگوریتم (یعنی تأثیر آن بر گروه‌های مختلف) را نیز ارزیابی کند. انتخاب معیار بهینه‌سازی (که ممکن است ترکیبی از دقت و عدالت باشد) یک تصمیم اخلاقی است که آماردان نمی‌تواند آن را صرفاً به کارفرما یا ناظران بیرونی واگذار کند.

ارزیابی و ممیزی: پس از استقرار سامانه، آماردان مسئول ممیزی مستمر عملکرد آن از نظر دقت، عدالت و عدم سوگیری است. این ممیزی‌ها باید مستقل و شفاف باشند.

در پاسخ به این چالش‌ها، حوزه‌های میان‌رشته‌ای جدیدی در سال‌های اخیر شکل گرفته‌اند. «یادگیری ماشین مسئول»^۲ و «اخلاق الگوریتم»^۳ تلاش می‌کنند تا روش‌های آماری و یادگیری ماشین را با ارزش‌های اجتماعی همچون عدالت، شفافیت، پاسخ‌گویی و حریم خصوصی پیوند دهند [۴۹]. این حوزه‌ها با پرسش‌های اساسی دست‌وپنجه نرم می‌کنند که پاسخ به آن‌ها نیازمند همکاری تنگاتنگ آماردانان، دانشمندان علوم کامپیوتر، حقوق‌دانان، فیلسوفان اخلاق و جوامع متأثر از الگوریتم‌هاست. برخی از این پرسش‌ها عبارت‌اند از:

تفسیرپذیری: آیا الگوریتم‌ها باید «تفسیرپذیر»^۴ باشند یا دقت بالای مدل‌های «جعبه سیاه»^۵ مجاز است؟ مدل‌های تفسیرپذیر، نظیر درخت تصمیم یا رگرسیون خطی، امکان درک نحوه تصمیم‌گیری یا پیش‌بینی خود را برای کاربر فراهم می‌کنند. در مقابل، مدل‌های جعبه سیاه، مانند شبکه‌های عصبی عمیق، فرایند داخلی خود را شفاف نمی‌سازند و کاربر نمی‌تواند به‌سادگی مسیر دستیابی به خروجی را دنبال کند. این مدل‌ها که اغلب از دقت بالایی برخوردارند، به دلیل عدم شفافیت، به‌ویژه در کاربردهای حساس همچون پزشکی یا قضایی، با چالش‌های اخلاقی جدی روبه‌رو هستند.

پاسخگویی: چه کسی در قبال سوگیری‌ها یا تولید محتوای

حامل سوگیری‌های انسانی باشند. به عنوان مثال، اگر به تصاویر پزشکی، برچسب «بیمار» یا «سالم» توسط پزشکانی زده شود که نسبت به برخی گروه‌های نژادی تشخیص‌های متفاوتی دارند، الگوریتم نیز همان تفاوت‌های تشخیصی را بازتولید می‌کند.

سوگیری در حجم نمونه‌های زیرگروه‌ها: اگر برخی از گروه‌های جمعیتی در داده‌های آموزشی نمونه کوچکی داشته باشند، الگوریتم پیش‌بینی‌هایی با دقت کم برای آن‌ها تولید می‌کند. این وضعیت به‌ویژه در گروه‌های اقلیت یا به‌حاشیه‌رانده شده مشاهده می‌شود.

۳.۶. عدالت الگوریتمی

یکی از چالش‌های اصلی اخلاق آماری در عصر داده‌های بزرگ، تعارض میان معیارهای سنتی دقت و معیارهای عدالت است. برای «عدالت الگوریتمی»^۱ تعریف‌های مختلفی از عدالت ارائه شده است که هر یک پیامدهای متفاوتی برای طراحی سامانه‌های آماری دارند [۵۹]:

عدالت به عنوان برابری آماری: الگوریتم باید پیش‌بینی‌های مثبت را با نسبتی یکسان در گروه‌های مختلف جمعیتی (مثلاً نژادی یا جنسیتی) تخصیص دهد. ایراد این تعریف آن است که اگر نرخ پایه پدیده در گروه‌ها متفاوت باشد (مثلاً نرخ واقعی وقوع جرم در دو گروه متفاوت باشد)، برابری آماری منجر به پیش‌بینی‌های نادقیق می‌شود.

عدالت به عنوان برابری نرخ خطا: لازم است الگوریتم به گونه‌ای طراحی شود که نرخ خطای مثبت و خطای منفی در گروه‌های گوناگون جمعیتی به یکسان باشد. این تعریف که درواکنش به نقدهای وارد بر برابری آماری شکل گرفت، به مراتب سخت‌گیرانه‌تر است و اغلب با دقت آماری در تعارض قرار می‌گیرد.

نکته حائز اهمیت آن است که این تعریف‌های مختلف از عدالت معمولاً با یکدیگر و با معیار دقت کلی ناسازگار هستند. به عبارت دیگر، در بسیاری از مسائل، نمی‌توان هم‌زمان دقت کلی را بیشینه کرد و همگی تعریف‌های عدالت را برآورده ساخت [۵۹]. این ناسازگاری، آماردانان را با پرسشی اساسی روبه‌رو می‌کند: کدام معیار را باید اولویت دهیم؟ پاسخ به این پرسش صرفاً فنی نیست، بلکه یک تصمیم اخلاقی و سیاسی است که باید با مشارکت ذی‌نفعان (از جمله گروه‌های آسیب‌پذیر) اتخاذ شود.

۴-۶. بازتعریف نقش آماردان

چالش‌های عصر داده‌های بزرگ، نقش و مسئولیت آماردان را دگرگون ساخته است. اگر در گذشته آماردان عمدتاً به عنوان تحلیلگر فنی در انتهای زنجیره پژوهش قرار داشت، امروزه از او انتظار می‌رود در

1. Algorithmic fairness
2. Responsible Machine Learning
3. Algorithmic ethics

4. Interpretable
5. Black-box

پزشکی مانند کامل الصناعات الطبیعة، فردوس الحکمة، الحاوی فی الطب (اثر معروف رازی)، ادب الطیب، طب روحانی، سرالطب، محنة الطیب و خواص التلامیذ فصل‌ها یا بخش‌هایی مستقل به اخلاق پزشکی اختصاص یافته است. این سنت غنی اخلاق حرفه‌ای در پزشکی ایرانی-اسلامی، زمینه‌ای فرهنگی برای توجه به اخلاق در سایر حوزه‌های علمی فراهم آورد [۱۸]. در همین سنت فلسفی-اخلاقی، کتاب «اخلاق ناصری» تألیف خواجه نصیرالدین طوسی (درگذشته ۶۷۲ قمری) رویداد مهمی در تدوین نظام‌مند اخلاق حرفه‌ای در تمدن ایرانی به شمار می‌رود. خواجه‌نصیر که در دوران اقامت اجباری خود در قهستان و به درخواست حاکم آن دیار، این کتاب را به زبان فارسی نگاشت، حکمت را به دو شاخه نظری و عملی تقسیم می‌کند و حکمت عملی را شامل سه بخش تهذیب اخلاق، تدبیر منزل و سیاست مدن می‌داند. او فضایل اخلاقی را در چهار جنس حکمت، شجاعت، عفت و عدالت خلاصه می‌کند و عدالت را به معنای مساوات و اعتدال تعریف می‌نماید. این مفهوم از عدالت به‌عنوان مساوات، پیشینه‌ای کهن در اندیشه اخلاقی ایران دارد که می‌تواند با اصول مدرن عدالت آماری (مانند پرهیز از سوگیری، انصاف در تحلیل و عدالت الگوریتمی) پیوند داده شود. همچنین تأکید خواجه‌نصیر بر رعایت حقوق دیگران و مسئولیت اخلاقی انسان در قبال اعمال خویش، می‌تواند زمینه‌ای فرهنگی برای شکل‌گیری خودآگاهی اخلاقی در میان دانشوران ایرانی فراهم آورد.

۷-۲. قوانین و شیوه‌نامه‌های اخلاقی

اگرچه آمار رسمی در ایران سابقه‌ای دیرینه دارد [۱۷] و سرشماری‌های مدرن از اواسط قرن چهاردهم شمسی به شکلی نظام‌مند اجرا شده‌اند [۱۶]، اما ملاحظات اخلاقی صریح و مدون در کارهای آماری، تا سال‌های اخیر حضوری آشکار در اسناد بالادستی کشور نداشته است. نخستین گام‌های نظام‌مند برای نهادینه‌سازی اخلاق در پژوهش در ایران، در حوزه علوم پزشکی برداشته شد. در سال ۱۳۷۰، آیین‌نامه اخلاق در پژوهش به تصویب رسید. در سال ۱۳۷۷، کمیته کشوری اخلاق تشکیل شد و در سال ۱۳۸۱، مرکز تحقیقات اخلاق و تاریخ پزشکی با هدف نظام‌مند کردن پژوهش‌های همه‌گیرشناسی و کارآزمایی‌های بالینی تأسیس گردید. از آنجا که این مطالعات به شدت به روش‌های آماری وابسته هستند، فعالیت‌های این مرکز به طور غیرمستقیم به نهادینه‌سازی ملاحظات اخلاقی در کارهای آماری پژوهش‌ها کمک کرد [۷]. در کنار تحولات پزشکی، قوانین بالادستی متعددی چارچوب حقوقی حفاظت از مالکیت فکری و اصالت علمی را در ایران شکل داده‌اند: قانون حمایت حقوق مؤلفان و مصنفان و هنرمندان (۱۳۴۸)، قانون ترجمه و تکثیر کتب و نشریات و آثار صوتی (۱۳۵۲)، قانون حمایت از حقوق پدیدآورندگان نرم‌افزارهای رایانه‌ای (۱۳۷۹)، و قانون پیشگیری و مقابله با تقلب در تهیه آثار علمی (۱۳۹۶) [۲، ۵، ۱۰، ۱۲]. قانون

نادرست، توهین آمیز یا تبعیض آمیز توسط این مدل‌ها و آسیب‌های ناشی از تصمیمات الگوریتمی پاسخ‌گوست؟ توسعه‌دهنده الگوریتم؟ صاحب داده؟ کاربر نهایی؟

حریم خصوصی: چگونه می‌توان حریم خصوصی را در عصر داده‌های بزرگ حفظ کرد بدون آنکه توانایی پژوهش‌های مفید اجتماعی را از دست بدهیم؟ آیا می‌توان به مدل‌هایی که بر روی میلیاردها داده استخراج شده از اینترنت (بدون آگاهی و رضایت صاحبان آن‌ها) آموزش دیده است، رضایت آگاهانه نسبت داد؟

مالکیت فکری: هم‌زمان با ظهور مدل‌های زبانی هوش مصنوعی مولد، پرسش‌های تازه‌ای درباره اصالت داده‌ها، مالکیت فکری، و مسئولیت در قبال محتوای تولیدشده توسط ماشین مطرح شده است. مدل‌های زبانی بزرگ بر روی میلیاردها متن، تصویر و اثر هنری از اینترنت آموزش می‌بینند که بسیاری از آن‌ها دارای حق مؤلف (حق نشر) هستند. آیا استفاده از این داده‌ها بدون اجازه صاحبان آن، نقض حق مؤلف محسوب می‌شود؟ اگر یک مدل زبانی، متنی تولید کند که شباهت زیادی به آثار یک نویسنده یا هنرمند خاص داشته باشد، آیا این خروجی، اثری اصیل و متعلق به توسعه‌دهنده مدل است، یا نقض حق مؤلف به شمار می‌رود؟ چه کسی صاحب حقوق معنوی این آثار است؟

این پرسش‌ها نشان می‌دهد که اخلاق آماری در آغاز مرحله‌ای تازه است؛ مرحله‌ای که در آن مرزهای سنتی «اخلاق در آمار» و «عدالت اجتماعی» از میان برمی‌خیزد. آماردان امروز دیگر تحلیلگری بی‌طرف نیست، بلکه ناگزیر به تأمل در نسبت داده، قدرت و عدالت اجتماعی است.

۷. اخلاق آماری در ایران

پرسش اساسی برای جامعه آماری ایران این است: تاریخچه توجه به ملاحظات اخلاقی در کارهای آماری در ایران چگونه بوده است؟ آیا می‌توان از اخلاق آماری در ایران به‌عنوان یک حوزه منضبط و دارای تاریخچه مستند سخن گفت؟ پاسخ به این پرسش، اگر با صداقت تاریخی داده شود، آمیخته از غیبت‌ها، ضرورت‌ها و فرصت‌هایی است که در ادامه به آنها می‌پردازیم.

۷-۱. ریشه‌های کهن

توجه به قوانین و مقررات اخلاقی در پژوهش (به‌ویژه در حوزه پزشکی) در ایران پیشینه‌ای دیرینه دارد. در ایران باستان، در وندیداد (یکی از بخش‌های اوستا) و نیز در قوانین حمورابی (که در میان‌رودان باستان رواج داشت و بر بخش‌هایی از ایران نیز تأثیر گذاشته بود)، به اخلاق پزشکی و حتی مجازات سرپیچی از آن اشاره شده است. این متون کهن نشان می‌دهند که ایده مسئولیت حرفه‌ای و پاسخگویی در قبال خطاهای علمی پیشینه‌ای چندین هزارساله در تمدن ایرانی دارد. پس از اسلام نیز، در کتب معتبر

انتشار و دسترسی آزاد به اطلاعات (۱۳۸۸) و قانون مدیریت داده‌ها و اطلاعات ملی (۱۴۰۱) نیز گام‌های مهمی در جهت ساماندهی به داده‌های ملی و شخصی برداشته‌اند [۶، ۳]. راهنمای کشوری اخلاق در انتشار آثار پژوهشی نیز در سال ۱۳۹۷ توسط کمیته ملی اخلاق در پژوهش به تصویب رسیده است [۹].

در حوزه‌های تخصصی دیگر نیز اسناد اخلاقی تدوین شده است: آئین رفتار حرفه‌ای برای حسابداران (۱۳۹۶)، نظام‌نامه رفتار حرفه‌ای اخلاقی در مهندسی ساختمان (۱۳۹۵) و نظام‌نامه اخلاق حرفه‌ای سازمان نظام روان‌شناسی (۱۳۸۶) [۱، ۴، ۸]. همچنین در حوزه هوش مصنوعی، راهنمای استفاده از ابزارهای هوش مصنوعی در پژوهش (بنیاد ملی علم، ۱۴۰۴) و آداب نامه استفاده از ابزار هوش مصنوعی مولد (دانشگاه تهران، ۱۴۰۴)، سند ملی هوش مصنوعی جمهوری اسلامی ایران (شورای عالی انقلاب فرهنگی، ۱۴۰۲) گام‌هایی در جهت تنظیم اخلاقی این فناوری‌ها محسوب می‌شوند [۱۳، ۱۴، ۱۵]. اما مهم‌ترین تحول در حوزه اخلاق آماری، تصویب شیوه‌نامه اخلاقی برای کارهای آماری در سال ۱۴۰۴ توسط انجمن آمار ایران است [۱۱]. این شیوه‌نامه که با الهام از شیوه‌نامه‌های اخلاق آماری انجمن‌های بین‌المللی و با توجه به بافت فرهنگی و حقوقی ایران تدوین شده، حوزه‌هایی مانند مسئولیت در قبال آزمودنی‌ها، حریم خصوصی داده‌ها، صداقت در تحلیل و تضاد منافع را پوشش می‌دهد. با این حال، صرف تصویب یک سند اخلاقی کافی نیست؛ چالش اصلی در اجرا و نظارت بر آن است. تجربه موفق کارگروه‌های اخلاق در پژوهش‌های پزشکی می‌تواند الگویی برای ایجاد ساختارهای مشابه در حوزه آمار باشد.

۷-۳ چالش‌های اخلاق آماری در ایران

فراتر از غیبت‌های نهادی، اخلاق آماری در ایران با چالش‌های بومی خاصی مواجه است:

چالش آموزش و فرهنگ‌سازی: آموزش دانشگاهی آمار در ایران از دهه ۱۳۴۰ تا امروز، عمدتاً تحت‌تأثیر رویکرد ریاضی‌وار و فنی بوده است. در این چارچوب، تاریخ، فلسفه و اخلاق آمار موضوعاتی حاشیه‌ای یا غیرضروری تلقی می‌شوند. مفاهیم اخلاقی به‌ندرت در دروس موجود مطرح می‌گردند. پژوهش‌های تاریخی و فلسفی درباره اخلاق آماری در ایران بسیار محدود است و این غیبت، به معنای فقدان خودآگاهی تاریخی در جامعه آماری ایران است. دانشجویان بیشتر با چگونگی تحلیل داده آشنا می‌شوند تا با چرایی انتخاب مسئله و پیامدهای اجتماعی کاربرد روش‌ها.

چالش داده‌های حساس: دسترسی به داده‌های حساس (مانند داده‌های پزشکی، مالی و رفتاری) با محدودیت‌های قانونی و امنیتی همراه است. «قانون انتشار و دسترسی آزاد به اطلاعات» [۶] و «قانون مدیریت داده‌ها و اطلاعات ملی» [۳] می‌توانند

چارچوبی برای دسترسی به داده‌ها ایجاد کرده، اما عدم شفافیت در فرایندهای دسترسی به داده‌های دولتی، پژوهشگران را با چالش دشواری روبه‌رو می‌کند: چگونه می‌توان پژوهش آماری معتبر انجام داد بدون آنکه به حریم خصوصی افراد تجاوز کرد یا با مقررات امنیتی در تضاد افتاد؟

چالش استانداردپذیری و تکرارپذیری: بحران تکرارپذیری در ایران با کمبود منابع و محدودیت دسترسی به داده‌های باکیفیت تشدید شده است. فشار برای افزایش تعداد مقالات منتشرشده می‌تواند به شیوع رفتارهای فرصت‌طلبانه (مشابه p -مقدار) منجر شود. با توجه به چالش‌های موجود، می‌توان دست‌کم چهار ضرورت را برای نهادینه‌سازی اخلاق آماری در ایران برشمرد:

توسعه و روزآمدسازی شیوه‌نامه اخلاقی: انجمن آمار ایران با تصویب شیوه‌نامه اخلاقی برای کارهای آماری در سال ۱۴۰۴ گام مهمی برداشته است. ضرورت دارد این شیوه‌نامه با مشارکت گسترده اعضای انجمن و با توجه به تحولات سریع فناوری (مانند هوش مصنوعی و داده‌های بزرگ) به طور پیوسته روزآمدسازی شود و سازوکارهای اجرایی و نظارتی مؤثری برای آن تدوین گردد. گنجاندن موضوع اخلاق آماری در برنامه درسی وزارت علوم و انجمن آمار ایران می‌تواند با افزودن دروسی با موضوع اخلاق آماری به برنامه درسی مقاطع مختلف، گامی اساسی در جهت ارتقای خودآگاهی تاریخی و اخلاقی دانشجویان بردارند. راهنمای استفاده از ابزارهای هوش مصنوعی در پژوهش [۱۳] می‌تواند الگویی برای تدوین راهنماهای مشابه در حوزه آمار باشد.

ایجاد سازوکارهای نظارت بر پژوهش‌های آماری: باید دامنه فعالیت کارگروه‌های اخلاق در پژوهش (که اکنون عمدتاً در علوم پزشکی فعال‌اند) به پژوهش‌های آماری در سایر حوزه‌ها نیز گسترش یابد. همچنین می‌توان سازوکارهای خاصی برای ممیزی مطالعات تکرارپذیر و بررسی شیوع p -مقدار در نشریات علمی ایران طراحی کرد.

تشویق به انجام پژوهش در حوزه اخلاق آماری: این مقاله خود دعوتی است برای جدی‌گرفتن پژوهش‌ها درباره اخلاق آماری و چالش‌های پیرامون آن و بهره‌گیری از تجربیات بین‌المللی برای بومی‌سازی اخلاق آماری در ایران. این موضوع نیازمند همکاری تاریخ‌نگاران علم، آماردانان و فیلسوفان اخلاق است. انجمن آمار ایران می‌تواند نقشی محوری در این مسیر ایفا کند. خبرنامه و نشریات انجمن به‌عنوان بستری برای گفتگوی غیر رسمی اما مستند میان اعضای انجمن، ظرفیت آن را دارد که:

- موضوع اخلاق آماری را از حاشیه به متن گفتگوهای حرفه‌ای بیاورد.
- تجربیات آماردانان ایرانی در مواجهه با چالش‌های اخلاقی را

گردآوری و مستند کند.

• زمینه‌ساز روزآمدسازی و توسعه شیوه‌نامه اخلاقی مصوب انجمن شود.

• پل ارتباطی میان جامعه آماری ایران و تجربیات بین‌المللی در زمینه اخلاق آماری ایجاد کند.

۸. نتیجه‌گیری

این مقاله با مروری بر تاریخ اخلاق آماری آغاز شد و با نگاهی به وضعیت ایران پایان یافت. فاصله میان نهادهای قدرتمند اخلاق‌پژوهش در کشورهای پیشرو و غیبت نسبی گفتگوی تخصصی در ایران، هم هشدار است و هم فرصت. هشدار از آن رو که غفلت از اخلاق آماری می‌تواند اعتبار پژوهش‌های علمی و اعتماد عمومی به آمار را خدشه‌دار کند. فرصت از آن رو که ایران می‌تواند با الهام از تجربیات بین‌المللی و تکیه بر ریشه‌های کهن اخلاق‌پژوهش در فرهنگ خود و تجربیات موفق کارگروه‌های اخلاق پزشکی، الگویی بومی و پیشرو در اخلاق آماری ارائه دهد. سیر تاریخی این مقاله روایتی از تحول تدریجی خودآگاهی اخلاقی جامعه آماردانان بود. این روایت چند درس اساسی برای آینده دارد

• روش‌شناسی صحیح و دقت فنی نمی‌تواند فقدان اخلاق در طراحی پژوهش یا کاربرد نتایج را جبران کند. اتکا به وجدان فردی برای جلوگیری از سوءاستفاده‌های نظام‌مند کافی نیست.

• اخلاق آماری نیازمند نهادهای بازدارنده است.

• فشارهای نهادی می‌توانند به فساد تدریجی رویه‌ها منجر شوند.

منابع

• عدالت، دقت را به چالش می‌کشد. آماردانی که تنها دقت کلی را پیشینه می‌کند، ممکن است ناآگاهانه به بازتولید نابرابری‌ها کمک کند.

در افق پیشرو، چالش‌های تازه‌ای در حوزه هوش مصنوعی مولد (پرسش از رضایت آگاهانه و مسئولیت در قبال محتوای تولیدشده)، تعارض حریم خصوصی با منفعت عمومی (تجربه کووید-۱۹)، و حکمرانی داده (نحوه جلوگیری از استثمار و تبعیض توسط شرکت‌های بزرگ فناوری) وجود دارد. تاریخ اخلاق آماری ما را به بازاندیشی در پرسشی اساسی فرامی‌خواند: آمار صرفاً مجموعه‌ای از روش‌های کمی است، یا حوزه‌ای از دانش با بارهای اجتماعی و اخلاقی عمیق؟ پاسخ دوم (که شواهد تاریخی این مقاله به نفع آن انباشته شده) اخلاق آماری را به عرصه‌ای گسترده بدل می‌کند که در آن پرسش از قدرت، نابرابری، عدالت و مسئولیت اجتماعی در مرکز توجه قرار می‌گیرد.

قدردانی و تشکر

این مقاله به پیشنهاد کارگروه اخلاق انجمن آمار ایران توسط نویسنده تهیه شده است. از نظرات ارزشمند اعضای کارگروه، آقایان دکتر محمد آرشی، دکتر علی رجالی، دکتر محمدرضا فقیهی، دکتر رحیم محمودوند و دکتر حمید نیلی ثانی در تدوین مقاله قدردانی می‌کنم. همچنین از آقای دکتر علی فرازمند سردبیر محترم نامه علوم پایه و آقای دکتر محمدقاسم وحیدی اصل که نسخه اولیه مقاله را مطالعه کردند و نظرات ارزشمندی ارائه نمودند، تشکر می‌کنم.

۱. آئین رفتار حرفه‌ای برای حسابداران (۱۳۹۶). جامعه حسابداران رسمی ایران.
۲. قانون حمایت حقوق مؤلفان و مصنفان و هنرمندان (۱۳۴۸). مصوب مجلس شورای ملی.
۳. قانون مدیریت داده‌ها و اطلاعات ملی (۱۴۰۱). مصوب مجلس شورای اسلامی.
۴. نظام‌نامه رفتار حرفه‌ای اخلاقی در مهندسی ساختمان (۱۳۹۵). مصوب وزارت راه و شهرسازی.
۵. قانون پیشگیری و مقابله با تقلب در تهیه آثار علمی (۱۳۹۶). مصوب مجلس شورای اسلامی.
۶. قانون انتشار و دسترسی آزاد به اطلاعات (۱۳۸۸). مصوب مجلس شورای اسلامی.
۷. قوانین، دستورالعمل‌ها و راهنماهای اخلاق در پژوهش‌های پزشکی (۱۴۰۱). مصوب وزارت بهداشت و درمان و آموزش پزشکی.
۸. نظام‌نامه اخلاق حرفه‌ای سازمان نظام روان‌شناسی (۱۳۸۶). مصوب سازمان نظام روان‌شناسی و مشاوره.
۹. راهنمای کشوری اخلاق در انتشار آثار پژوهشی (۱۳۹۷). مصوب کمیته ملی اخلاق در پژوهش.
۱۰. قانون حمایت از حقوق پدیدآورندگان نرم‌افزارهای رایانه‌ای (۱۳۷۹). مصوب مجلس شورای اسلامی.
۱۱. شیوه‌نامه اخلاقی برای کارهای آماری (۱۴۰۴). مصوب هیئت‌مدیره انجمن آمار ایران.
۱۲. قانون ترجمه و تکتیر کتب و نشریات و آثار صوتی (۱۳۵۲). مصوب مجلس شورای ملی.
۱۳. راهنمای استفاده از ابزارهای هوش مصنوعی در پژوهش (۱۴۰۴). مصوب بنیاد ملی علم.
۱۴. آداب نامه استفاده از ابزار هوش مصنوعی مولد در تهیه مستندات علمی دانشجویان در دانشگاه تهران (۱۴۰۴). مصوب دانشگاه تهران.
۱۵. سند ملی هوش مصنوعی جمهوری اسلامی ایران (۱۴۰۲). مصوب شورای عالی انقلاب فرهنگی.
۱۶. وحیدی اصل، محمدقاسم (۱۴۰۳). تاریخ آمار، انتشارات مبتکران.

۱۷. تنکابنی، حمید (۱۳۸۳). *درآمدی بر دیوان سالاری در ایران*، انتشارات علمی و فرهنگی، تهران.
۱۸. نیک روان فرد، نازیلا (۱۴۰۲). تاریخچه اخلاق در پژوهش معاصر، *دایره المعارف اخلاق پزشکی اسلامی*، صص ۲۷-۱.
19. Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016). *Machine bias*. ProPublica.
20. Annas, G. J. and Grodin, M. A. (Eds.). (1992). *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*. Oxford University Press.
21. ASA (American Statistical Association). (1981). *Ethical Guidelines for Statistical Practice*.
22. ASA (American Statistical Association). (2016). ASA statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2),129–133.
23. ASA (American Statistical Association). (2018). *Ethical Guidelines for Statistical Practice* (Revised ed.).
24. Bailar, J. C. and Bailar, M. (1983). Ethics in statistics. *The American Statistician*, 37(4),283–287.
25. Bankert, E. A. and Amdur, R. J. (2006). Institutional Review Board: Management and Function. Jones and Bartlett Pub-lishers.
26. Barocas, S. and Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3),671–732.
27. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research. U.S. Government Printing Office.
28. Berger, R. L. (1990). Nazi science: The Dachau hypothermia experiments. *New England Journal of Medicine*, 322(20),4410–5341;
29. Bommasani, R., et al. (2021). On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258.
30. Box, J. F. (1978). R. A. Fisher: The Life of a Scientist. Wiley.
31. Brandt, A. M. (1978). Racism and research: The case of the Tuskegee Syphilis Study. *Hastings Center Report*, 8(6),12–92
32. Brown, T. H. (1952). The statistician and his conscience. *The American Statistician*, 6(5),14–18..
33. Buck v. Bell, 274 U.S. 200(1927).
34. Burleigh, M. (1994). *Death and Deliverance: Euthanasia in Germany c. 1900-1945*. Cambridge: Cambridge University Press.
35. Court, A. T. (1952). Standards of statistical conduct in business and in government. *The American Statistician*, 6(5),6–41
36. Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25(1),7–29..
37. Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters.
38. Davenport, C. B. (1911). *Heredity in Relation to Eugenics*. New York: Henry Holt & Company.
39. Deming, W. E. (1985). Ethics and statistics. *The American Statistician*, 39(4),253–254.
40. Desrosières, A. (1998). *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press
41. Dodson, S. and Bartley, J. (2024). (Disrupting) continuities between eugenics and statistics: A critical study of regression analysis. *Proceedings of the Association for Information Science and Technology*, 61(1),892–895..
42. Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3-4),211–407.
43. Ebert, F. W. (2001). Johann Peter Süßmilch, Pfarrer, Universalwissenschaftler und Vater der deutschen Statistik. *Statistische Geschichten, Daten Analysen*, IV/2001. Retrieved from www.statistik-berlin-brandenburg.de
44. Ellenberg, J. H. (1983). Ethical guidelines for statistical practice: A historical perspective. *The American Statistician*, 37(1),1–4.

45. European Union. (2016). General Data Protection Regulation (GDPR). Official Journal of the European Union, 119, 1–88.
46. Fellegi, I. P. (1996). The independence of official statistics. Statistical Journal of the United Nations Economic Commission for Europe, 13(3), 203–212.
47. Fienberg, S. E. (1994). Ethics and statistics. International Statistical Review, 62(1), 1–6.
48. Fisher, R. A. (1924). The biometrical study of heredity. Eugenics Review, 16(3), 189–210.
49. Floridi, L., et al. (2018). AI⁴People: An ethical framework for a good AI society. Minds and Machines, 28(4), 689–707.
50. Freeman, W. W. K. (1952). Discussion of The statistician and his conscience by T. H. Brown. The American Statistician, 6(5), 18–20.
51. Galton, F. (1869). Hereditary Genius: An Inquiry Into its Laws and Consequences. Macmillan.
52. Gelman, A. and Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no fishing expedition or p-hacking and the research hypothesis was posited ahead of time. Department of Statistics, Columbia University.
53. Gigerenzer, G. (2004). Mindless Statistics. The Journal of Socio-Economics, 33(5), 587–606.
54. Hacking, I. (1990). The Taming of Chance. Cambridge, UK: Cambridge University Press.
55. Heller, J. (1972). Syphilis victims in U.S. study went untreated for 40 years. Associated Press.
56. Ioannidis, J. P. (2005). Why most published research findings are false. PLoS Medicine, 2(8), e124.
57. ISI (International Statistical Institute). (1985). Declaration On Professional Ethics.
58. Jones, J. H. (1993). Bad Blood: The Tuskegee Syphilis Experiment (New and expanded ed.). Free Press.
59. Kleinberg, J., Mullainathan, S. and Raghavan, M. (2017). Inherent trade-offs in the fair determination of risk scores. th Innovations in Theoretical Computer Science Conference.
60. Kühl, S. (1994). The Nazi connection: Eugenics, American Racism, and German National Socialism. New York: Oxford University Press.
61. Laughlin, H. H. (1922). Eugenical Sterilization In the United States. Chicago: Psychopathic Laboratory of the Municipal Court of Chicago.
62. Lombardo, P. A. (2008). Three Generations, no Imbeciles: Eugenics, the Supreme Court, and Buck v. Bell. Baltimore: Johns Hopkins University Press.
63. MacKenzie, D. (1981). Statistics in Britain, 1865–1930: The social construction of scientific knowledge. Edinburgh University Press.
64. Malthus, T. R. (1798). An Essay on the Principle of Population as it Affects the Future Improvement of Society. London.
65. Mittelstadt, B. D. and Floridi, L. (Eds.). (2016). The Ethics of Biomedical Big data. Springer.
66. Morton, J. E. (1952). Standards of statistical conduct in business and government. The American Statistician, 6(5), 6–7.
67. Munafò, M. R., et al. (2017). A manifesto for reproducible science. Nature Human Behaviour, 1(1), 1–9.
68. Nosek, B. A., et al. (2015). Promoting an open research culture. Science, 348(6242), 1422–1425.
69. Nuremberg Code. (1947). Permissible medical experiments. Trials of War Criminals before the Nuremberg Military Tribunals, 2, 181–182.
70. O’Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishers.
71. Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349 (6251) aac4716.
72. Pearson, K. (1901). National life from the standpoint of science. *Nature*, 64(1661), 441–442.
73. Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334(6060), 1226–1227.

74. Porter, T. M. (1995). *Trust In Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press.
75. Proctor, R. N. (1988). *Racial Hygiene: Medicine under the Nazis*. Harvard University Press.
76. Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641.
77. Simmons, J. P., Nelson, L. D. and Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366.
78. Stigler, S. M. (2000). *The History of Statistics: The measurement of uncertainty before 1900*. Cambridge, MA: Harvard University Press. (Original work published 1986)
79. Wheatcroft, S. G. (1990). More light on the scale of repression and excess mortality in the Soviet Union in the 1930s. *Soviet Studies*, 42(2), 355–367.
80. World Health Organization. (2020). *Ethical Considerations to Guide the Use of Digital Proximity Tracking Technologies for COVID-19 Contact Tracing*. WHO.