پیک ریاضی

جلد دوم، شماره سوم پائیز ۱۳۶۶

مقالهٔ ۱۷۷۴ لاپلاس دربارهٔ احتمال وارون

نوشتهٔ : استفان م. استایگلر

ترجمهٔ : محمدحسین علامت ساز

چکیده.

اولین مقاله عمده لاپلاس در آمار ریاضی در سال ۱۷۷۴ منتشر شد. قطعاً، مقاله مزبور پرنفوذترین مقاله ای است که در این زمینه تا قبل از سال ۱۸۰۰ ارائه شده و در ضمن اولین مطالعه گسترده احتمال وارون وکاربردهای آن در برآورد پارامتر و جمله ای و پارامتر محلی است.

مقدمه.

تاریخ آمار ریاضی تا قبل از سال ۱۸۰۰ چند رسالهٔ برجسته را مانند هنر برآورد کردن از یاکوب برنولی (۱۷۱۳) و آموزهٔ شانسها از دموآور (۱۷۱۸، چاپ دوم آن ۱۷۳۸) و تعداد کمتری مقالات مهم در مجلات ادواری را شامل می‌شود. شاید پرنفوذترین آنها از نوع دوم مقالهٔ لاپلاس دربارهٔ احتمال وارون باشد که در سال ۱۷۷۴ چاپ شد. در آن زمان لاپلاس ۲۵ سال داشت و این اولین کار قابل توجه او در آمار ریاضی بود. زمانی که در سال ۱۷۷۲ — این کار را شروع کرد او در دوران خلاقیت جدی اکتشافات علمی بود. او به طور

همزمان پیشرفتهای قابل توجهی درریاضیات ونجوم ریاضی داشت واین مقالهٔ اوکه فوران ناگهانی ایده ها است اثری محونشدنی درآمار بجای گذاشت. درهمین یک مقاله می‌توانیم ریشه های نظریه تصمیم‌گیری نوین، استنباط بیزی با پارامترهای مجهول وتقریب مجانبی توزیع پسین را شناسایی کنیم.

ارائه ارزیابی کاملی ازسهم لاپلاس درزمینه ای تاریخی ازوسع این مقاله خارج است. با وجوداین وعلیرغم این حقیقت که مقاله روشن وروان جلوه می‌کند (حتی پس ازدوقرن مانند یک کارمعاصر بنظرمی‌رسد)، طرح مختصری ازآن درواژه های مدرن برای اغلب خوانندگان مفیدخواهد بود.

لاپلاس پس ازاشاره به منابعی ازکارهای خود ومقالهٔ لاگرانژ درمورد حل معادلات تفاضلی (بخصوص آنهایی که درنظریه احتمال روی می‌دهند) دربخش دوم به مسائل استنباط می‌پردازد که محوراصلی توجه مقاله است. او اولین هدف خودرا بروشنی تعیین یک احتمال دوجمله ای مجهول معرفی می‌کند که درآن از $p+q$ آزمایش، p تا به برد و q تا باخت منجر شده اند. او اصلی را بیان می‌کند که امروزه ما آن را معادل قضیه بیز با علتهای پیشین متساوی الاحتمال می‌شناسیم. اگر F یک پیشامد لاپلاس با شد با n علت $\theta_1, \theta_2, \ldots, $ و θ_n آنگاه "اصل" اصولی او عبارت است از

$$\frac{P(\theta_i|F)}{P(\theta_j|F)} = \frac{P(F|\theta_i)}{P(F|\theta_j)}$$

و

$$P(\theta_i|F) = \frac{P(F|\theta_i)}{\sum_{j=1}^{n} P(F|\theta_j)}$$

بعضی ازدلایلی که بطورمعقول ما راازناآگاهی لاپلاس ازکارقبلی بیز مطمئن می‌سازدرا می‌توان دراستایگلر (۱۹۷۸) یافت .

درواقع ، یکی ازقانع کننده‌ترین دلایل ، روش بسیارمتفاوتی است که‌به‌کاربرده‌شده‌است : درجایی که‌بیزبحثی قوی درموردچگونگی امکان پذیرش یک توزیع پیشین یکنواخت ارائه‌می‌دهد (استایگلر ۱۹۸۲) ، لاپلاس این نتیجه‌را به‌عنوان فرضی شهودی وواضح فرض می‌کند . لاپلاس یک کاربردساده‌ازاصل خودرادرحالت نمونه‌گیری فوق هندسی ازیکی ازدوظرف ممکن A و B درنظرگرفته‌ومقدار

K=P {ظرف A | کشیدن f تا سفیدو h تا سیاه}

را محاسبه‌کرده ، بدست می‌آورد .

P{ظرف A | کشیدن f تا سفیدو h تا سیاه }=k/(k+k')

(که‌درآن { ظرف B |کشیدن f تا سفیدو h تا سیاه } K'=P) که‌این نتیجه‌ء صحیحی است اگر A و B پیشین های با احتمال مساوی باشند .

دربخش سوم ، لاپلاس این اصل رادرحالت نمونه‌گیری دوجمله‌ای با یک توزیع پیشین یکنواخت برای پارامترآن به‌کارمی‌برد . درمساله اول ، اوتوزیع پسین بتا را برای پارامتردوجمله‌ای x وهمچنین توزیع پیشگوی لاپلاس (E) را برای m+n آزمایش آینده به‌شرط آنکه p+q آزمایش رخ داده‌باشد ، ابتدا برای m=1 و n=0 وسپس برای حالت عمومی محاسبه می‌کند . دراین موردا وفرمول استرلینگ را به‌صورتی که‌ازیک رساله‌ء لئونهارد اویلربه‌دست آورده‌جهت محاسبه‌ء تقریبی نمونه‌های بزرگ به‌کار می‌بردومحاسبات مجانبی را به‌دوصورت مختلف انجام می‌دهد . ابتدا فرض می‌کندکه p+q بزرگ است (درصورتی که m+n بزرگ نیست) ، ونتیجه

می‌گیرد که توزیع پیشگو تقریبا" مساوی حالتی است که در آن x برابر نسبت نمونه یعنی $\frac{p}{p+q}$ اختیار شود. سپس نشان می‌دهد که اگر m+n نیز بزرگ فرض شود، مثلا" $m+n=p+q$، نتیجه‌ای متفاوت به دست می‌آید. در باقیمانده این بخش، آنالیز توزیع پسین برای نمونه‌های بزرگ مورد نظر گرفته می‌شود. او قضیه‌ای بیان می‌کند (با کمال تعجب به صورت $\varepsilon-\delta$ ی امروزی) که در آن سازگاری پسین فراوانی نسبی موفقیت‌ها را ادعا می‌کند ولی اثبات او قویتر است و کاری بیشتر از آن انجام می‌دهد. آن هم روش خود لاپلاس را برای محاسبات تقریبی مجانبی انتگرالها بوسیلهٔ بسط تابع انتگرالگیری حول ماکزیمم آن در $x=p/(p+q)$ معرفی می‌نماید و هم بطور موثری نرمال بودن مجانبی توزیع پسین بتا را هنگام تقریب کردن احتمال پسین

$$E=P\left\{\left|x-\frac{p}{p+q}\right|<\omega \;\middle|\; p \text{ بلیط سفید و } q \text{ تا سیاه}\right\}$$

توسط یک انتگرال نرمال وحدگیری وقتیکه $p+q\to\infty$ با $\omega=(p+q)^{-\frac{1}{n}}$ و $2<n<3$ نمایش می‌دهد. چون این اثبات شامل چیزی است که ظاهرا" اولین محاسبهٔ انتگرال معین

$$2\int_0^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2\sigma^2}\right)dz=1$$

با $\sigma^2=pq/(p+q)^3$ باشد لاپلاس ممکن است اولین کسی بوده باشد که چگالی نرمال را انتگرالگیری کرده است. در سال ۱۸۰۹، گوس به این محاسبات بعنوان "قضیهٔ زیبای لاپلاس" اشاره می‌کند. این قسمت با یک تلاش بغرنج جهت تقریب کردن خطای انجام شده، در صورتی که $E=1$ فرض شود به اتمام می‌رسد.

بخش چهارم به مسالهٔ دوم، تحلیلی از مسالهٔ کلاسیک "نقاط" از دیدگاهی بیزی، می‌پردازد. اکنون ممکن است ما این مساله را به عنوان

پیداکردن امیدریاضی توزیع پیشگوی دقیق برای نمونه‌گیری دوجمله‌ای با توزیع پیشین یکنواخت توصیف کنیم.

دربخش پنجم، لاپلاس درجهت برآوردکردن یک پارامترمحلی برای حالت "سادهٔ" سه‌مشاهده‌ای حرکت می‌کند. ظاهرا" اوبا خواندن پانوشتی کوتاه درمقاله‌ای مروری ازژان برنولی سوم که در ۱۷۷۲ چاپ شده بود به وجدآمده بود: "مسالهٔ یافتن میانگین واقعی چندمشاهده که بندرت میانگین‌حسابی است بطورقابل ملاحظه‌ای موردعلاقه منجمین است." برنولی کارهایی ازبوسکوویچ ولامبرت راکه منتشرشده بودندوازلاگرانژودانیل برنولی راکه بعدا"درسالهای ۱۷۷۶ و۱۷۷۸ چاپ شدندذکرکرده بود. (تمام پانوشت با منابع دراستایگلر (۱۹۷۸ صفحهٔ ۲۴۸) دوباره چاپ گردیده است) لاپلاس بررسی این مساله (" مسالهٔ دوم") راباا ندیشیدن دربارهٔ طبیعت توزیع خطا ویافتن عبارت عمومی برای توزیع پسین چیزی که ما حالا آن راپارامترمحلی می‌خوانیم آغازمی‌کند. شکل ۱ سه‌مشاهدهٔ a, b و c را نشان می‌دهدوشکل ۲ توزیع خطا را. v مقدارواقعی پارامترمحلی‌و p و q بترتیب شکاف بین‌a و b و b و c رانمایش می‌دهند. حال ممکن است بنویسیم

$$q=x^{(3)}-x^{(2)}\ ,\ p=x^{(2)}-x^{(1)}\ ,\ c=x^{(3)},\ b=x^{(2)}\ ,\ a=x^{(1)}$$

وتوجه‌کنیم‌که p و q آماره‌های فرعی هستند. لاپلاس x رابه‌عنوان فاصلهٔ v تا a قرارمی‌دهد. شکل ۱ شرح ترسیمی توزیع پسین‌x مشروط بر داشتن p و q رابه‌دست می‌دهدیعنی منحنی HOL را. لاپلاس هردوی میانه پسین ("میانگین احتمال") ومقداری که امیدریاضی پسین زیان ("میانگین خطا") رامینیموم می‌کندبه‌عنوان برآوردپسین پارامترمحلی پیشنهادکرده واثبات می‌کند (باکمک شکل ۳) که اینها همواره مساوی اند.

این نتیجه (مشخص‌سازی میانهٔ پسین به‌عنوان بهینه‌برای تابع زیان

Fig. 1.

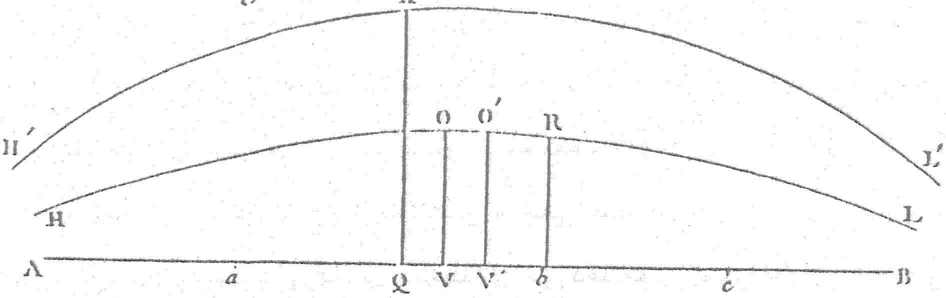


Fig. 2.

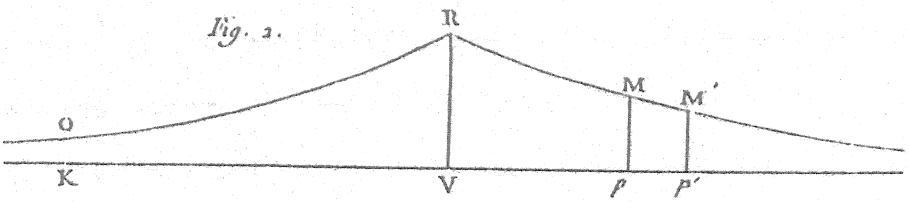
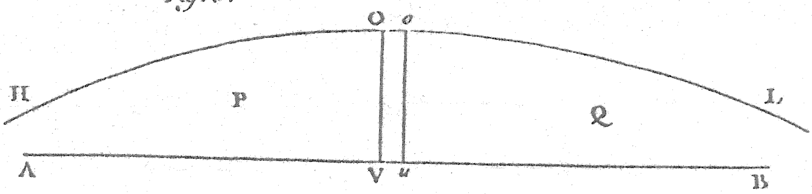


Fig. 3.



معینی)مطمئنا "یکی ازاولین نتایج شناسائی شده است که ما می‌توانیم واقعا "به آمار ریاضی متعلق بدانیم تا به نظریهٔ احتمال. سپس او به سوال تصویح کردن توزیع خطا بازگشته و بحثی درباب چگالی نمائی دوگانه

$$\psi(x) = \frac{1}{2} m \ e^{-m|x|}$$

ارائه می‌دهد. با فرض اینکه پارامتر مقیاس m معلوم است، او عبارتی صریح برای میانهٔ پسین می‌یابد و همان طور که ژان برنولی سوم گفته بود نشان می‌دهد که آن با میانگین حسابی که عموما "استفاده می‌شد متفاوت است.

بقیهٔ بخش پنجم مربوط به حالتی است که پارامتر مقیاس m مجهول است اما توزیع پیشین یکنواخت است. تحلیل او در این قسمت بغرنج و پیچیده است. این پیچیدگی به خاطر خطای ازنظرگریزنده‌ای که در یافتن یک عبارت صریح برای میانهٔ پسین انجام می‌دهد بیشتر می‌شود. گرههای اصلی مقاله در دو پاراگراف متوالی است که به دنبال محاسبهٔ درست توزیع پسین m آمده‌اند، پاراگرافهایی که با "پس، اگر ما توسط y نشان دهیم ...." شروع می‌شوند. چون ظاهرا "طبیعت و اهمیت این خطا از نظر سایر مفسرین به دور مانده است ارزشمند است که تفسیر مفصلی از آن در اینجا آورده شود. برای کمک به روشنی بحث، علامت f را به عنوان نماد عمومی برای چگالی به کار می‌بریم. پس مثلا"، f(x,m|p,q) نمایشگر چگالی شرطی x و m است به شرط آنکه p و q معین باشند. به یاد آورید که لاپلاس مساله‌اش را برحسب تصحیح، x، که در مشاهدهٔ اول، a، می‌سازد مطرح کرده و دنبال میانهٔ پسین x (که آن را با x_0 نشان خواهم داد) است. لاپلاس قبل از آن، به درستی در حالتی که m را معلوم فرض می‌کند، رابطه‌ای که میانهٔ پسین را معین می‌کند به صورت

$$\int_{-\infty}^{x_0} f(x,p,q|m)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} f(x,p,q|m)\,dx$$

به‌دست می‌آورد. چون این چگالی توام متناسب با چگالی شرطی $f(x|p,q,m)$ است با ثابت تناسب $\left[\int_{-\infty}^{\infty} f(x,p,q|m)\,dx\right]^{-1}$، برابری بالامعادل است با

$$\int_{-\infty}^{x_0} f(x|p,q,m)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} f(x|p,q,m)\,dx$$

دراینجا، وقتیکه m مجهول باشدتحت این فرض‌که m یک پیشین است‌که بطوریکنواخت روی $(0,\infty)$ توزیع شده، او $f(m|p,q)$ رابه‌صورت

$$f(m|p,q) \propto m^2\, e^{-m(p+q)}\left(1-\frac{1}{3}e^{-m/p}-\frac{1}{3}e^{-mq}\right)$$

محاسبه می‌کند. حال y اوعبارت است از $f(x|p,q,m)$ وظاهرا" لاپلاس قصددارد‌به‌صورت زیرعمل کند:

ما می‌خواهیم x_0 طوری باشدکه

$$\int_{-\infty}^{x_0} f(x|p,q)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} f(x|p,q)\,dx,$$

ویا معادل است با

$$\int_{0}^{\infty}\int_{-\infty}^{x_0} f(x,m|p,q)\,dx\,dm = \frac{1}{2}\int_{0}^{\infty}\int_{-\infty}^{\infty} f(x,m|p,q)\,dx\,dm,$$

ویا

$$\int_{0}^{\infty}\int_{-\infty}^{x_0} f(x|m,p,q)\,f(m|p,q)\,dx\,dm$$

$$= \frac{1}{2}\int_{0}^{\infty}\int_{-\infty}^{\infty} f(x|m,p,q)\,f(m|p,q)\,dx\,dm$$

حال اگر این آخرین برابری معادل می‌بود با

$$\int_0^\infty \int_{-\infty}^{x_0} f(x,p,q|m)\,f(m|p,q)\,dx\,dm = \frac{1}{2}\int_0^\infty \int_{-\infty}^{\infty} f(x,p,q|m)\,f(m|p,q)\,dx\ dm,$$

آنگاه ما واقعا "به حل لاپلاس می‌رسیدیم. ولی این هنگامی درست است که

$$f(x|m,p,q) \propto f(x,p,q|m)$$

و ثابت تناسب به m بستگی نداشته باشد که اینجا چنین نیست. درحقیقت، برای پیشین لاپلاس داریم:

$$f(x|m,p,q)=f(x,p,q|m)/f(p,q|m)$$

$$\propto f(x,p,q|m)/f(m|p,q),$$

بنابراین حل صحیح باید به صورت

$$\int_0^\infty \int_{-\infty}^{x_0} f(x,p,q|m)\,dx\,dm = \frac{1}{2}\int_0^\infty \int_{-\infty}^{\infty} f(x,p,q|m)\,dx\,dm\ ,$$

باشد، یک جفت انتگرال که حتی ساده‌تر از آنهایی است که لاپلاس درنظر گرفت. لاپلاس به خاطر خطایش مجبور به یافتن ریشهٔ یک معادلهٔ درجهٔ پانزدهم شد که به روش تکراری آن را حل کرد و جواب را به شکل یک جدول ارائه داد. جای تعجب نیست که چرا او فقط حالت سه مشاهده‌ای را درنظر می‌گیرد. خطای لاپلاس مهم است زیرا که روشن می‌سازد که چرا او توزیع های شرطی را فقط به عنوان تناسب پنداشته و نیز به روشن شدن این نکته که چگونه "اصل" اصولی چنین شهودی‌ در او برانگیخته است کمک می‌نماید (استایگلر، ۱۹۸۶، فصل ۳).

سرانجام، بخش ششم موثر بودن یک توزیع پیشین برای احتمال را دردو حالت کلاسیک پرتاب سکه و انداختن یک تاس کشف می‌کند. کار او در این بخش

با بخشهای پیش از این نظر متفاوت است که در اینجا یک پیشین غیریکنواخت را درنظر می‌گیرد. لاپلاس عملا" فرض صفر صریح " سالم بودن سکه" را بدین صورت غیرصریح می‌کندکه ابتدا احتمال را به جای $\frac{1}{2}$ برابر($\pi-1$) درنظر می‌گیرد وسپس فرض می‌کندکه π روی $[0,q^{-1}]$ بطوریکنواخت توزیع شده باشد. او متوجه می‌شودکه این نتایج برای پیشامدهای مرکب (بخصوص امید ریاضی توزیع پیشگو) می‌توانندبه نحوچشمگیری با حالتی که یک سکهٔ سالم درنظرگرفته می‌شود متفاوت باشند. تحلیل وی برای تاس، اگرچه پیچیده‌تر است ولی مشابه همین است. جالب توجه است که او تاس ناسالم را "تاس انگلیسی" می‌نامد. ای بساکه لندنیها لقب متفاوتی برای تاس مزبور داشته باشند. او با محاسبات صریحی برای حالت یک تاس ۳ وجهی مطلب را به پایان می‌رساند ونتایجی را استخراج می‌کندکه باید از اولین کاربردهای ایده‌های بیزی درحالت چندجمله‌ای باشند.

نفوذ این مقاله از اینجا بودکه ایده‌های "بیزی" برای اولین بار درسراسر دنیای ریاضیات پخش شد، درحالی که مقالهٔ خودبیز (بیز ۱۷۶۴) تا بعد از ۱۷۸۰ نادیده گرفته شده بود وتا قرن بیستم هیچ نقش مهمی درمباحث علمی ایفا نکرد. (استایگلر ۱۹۸۲). همچنین این مقالهٔ لاپلاس بودکه تکنیکهای ریاضی برای آنالیزمجانبی توزیعهای پسین راکه امروزه هنوز بکار گرفته می‌شوند معرفی کرد. ودر اینجاست که می‌توان اولین مثال برآورد بهینه را با محاسبه ومشخص نمودن یک برآوردکننده که اندازه مخصوص از امید ریاضی پسین زیان را مینیمم می‌کند یافت. اکنون پس از نزدیک به دو قرن، آمارشناسان نه تنها ریشه‌ها را در این شاهکار علمی شناسائی می‌کنند بلکه هنوزهم از آن چیزیاد می‌گیرند.

Stephen M. Stigler,

Laplace's 1774 Memoir on Inverse Probability, *Statistical Science*, 1986,Vol.1, No.3, 359-362.

BAYES, T. (1764). An essay towards solving a problem in the doctrine of chances. Philos. Trans. Roy. Soc.London 53 370-418.(Reprinted,with an introduction by George Barnard,in 1958 in Biometrika 45 293-315, and in 1970 in Studies in the History of Statistics and Probability (E.S. Pearson and M.G. Kendall,eds.) 131-153. Hafner,Darien,Conn.)

DALE, A. I. (1982).Bayes or Laplace? An examination of the origin and early applications of Bayes's theorem.Arch.History Exact Sci. 27 23-47.

GILLISPIE, C.C. (1981).Laplace,Pierre Simon.Dictionary of Scientific Biography 15,273-403. Scribner,New York.

LAGRANGE,J.L. (1759).Sur l'integration d'une equation differentielle a differences finies,qui contient la theorie des suites recurrentes.Miscellanea Philos.Math.Soc.Privatae Taurinenesis 1 33-42.

LAPLACE,P.S. (1774a).Memoire sur les suites recurro-recurrentes et sur leurs usages dans la theorie des hasards.Memoires de mathematique et de physique presentes a l'Academie royale des sciences, par divers savans,& lùs dans ses assemblees 6 353-371.(Reprinted in Laplace's Oeuvres completes 8 5-24.)

LAPLACE, P. S. (1774b).Memoire sur la probabilite des causes par les evenemens.Memoires de mathematique et de physique presentes a l'Academie royale des sciences,par divers savans, & lus dans ses assemblees 6 621-656.(Reprinted in Laplace's Oeuvres completes 8 27-65.)

LAPLACE, P. S. (1775).Memoire,sur les solutions particulieres des equations differentielles, & sur les inegalites seculaires des planetes. Histoire de l;Academie royale des sciences, annee 1772,premiere partie,avec les Memoires de mathematique et de physique,pour la meme annee,tires des registres de cette academie (published 1775) 343-377 (Reprinted in Laplace's Oeuvres completes 8 325-366.)

LAPLACE, P. S. (1776).Recherches,sur l'integration des equations differentielles aux difference finies,& sur leur usage dans la theorie des hasards.Memoires de mathematique et de physique presentes a l'Academie royale des sciences,par divers savans, & lus dans ses assemblees, 1773 (published 1776)37-162. (Reprinted in Laplace's Oeuvres completes 8 69-197,)

LAPLACE, P. S. (1886-1912). Oeuvres completes de Laplace. 14 volumes.Gauthier-Villars,Paris.

SHEYNIN,O. B. (1977). Laplace's theory of errors. Arch. History Exact Sci 17 1-61.

STIGLER, S. M. (1978).Laplace's early work: Chronology and citations.Isis 61 234-254.

STIGLER, S. M. (1982). Thomas Bayes's Bayesian inference.J.Roy. Statist.Soc.Ser.A 143 250-258.

STIGLER,S.M. (1986). The History of Statistics: The Measurement of Uncertainty Before 1900.Harvard Univ.Press,Cambridge, Mass.

TODHUNTER, I.(1865). A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace. Macmillan,London.(Reprinted by Chelsea,New York,in 1949, 1965.)