H. STEINHAUS (Wrocław)

PODSTAWY KONTROLI STATYSTYCZNEJ

Celem tej noty jest wyjaśnienie pewnej kontrowersji, która pojawia się w nauce o prawdopodobieństwach od samych początków tej dyscypliny w XVIII wieku, a którą kojarzymy z tak zwaną *regułą Bayesa*[1]). Sprawa *prawdopodobieństwa przyczyn* — taki jest termin techniczny — należy do teorii i może się wydać dziwne, że chcemy ją traktować tutaj zasadniczo na tle szczegółowego zagadnienia praktycznego, jakim jest statystyczna kontrola jakości. Bierze się to stąd, że kontrowersja dotyczy postępowania, a nie twierdzeń teoretycznych; nie istnieje ona w zakresie matematyki czystej. Nikt nie wątpi o poprawności wzoru Bayesa na tak zwane *prawdopodobieństwo a posteriori*, tak jak nie można nic zarzucić koncepcji *przedziału ufności* przyjętej przez Spławę-Neymana i jego szkołę. Ale w praktyce kontroli statystycznej ta koncepcja niemal zupełnie wyrugowała postępowanie oparte na obliczaniu prawdopodobieństw przyczyn. Wprawdzie w kołach przyrodników i lekarzy używa się często wyrazu „prawdopodobieństwo" w znaczeniu zakazanym przez nową szkołę, jednak nie świadczy to na rzecz starej doktryny, bo w tych przypadkach można zarówno napotkać błąd istotny polegający na nieznajomości klasycznego rachunku prawdopodobieństwa, jak błąd formalny polegający na nieznajomości terminologii.

---

[1]) Por. np., co mówi W. Feller (*An introduction to probability theory and its applications*, I (1950), str. 85): „Inżynier ma przy kontroli jakości do czynienia z jedną maszyną, a nie z nieskończoną populacją maszyn, z której tę jedną wybrano na chybił-trafił. Poradzono mu używać reguły Bayesa z tej racji, że jest ona logicznie do przyjęcia i odpowiada naszemu sposobowi myślenia. Plato używał podobnego argumentu dla udowodnienia istnienia Atlantydy, a filozofowie do wykazania niedorzeczności mechaniki Newtona... Współczesna teoria testów statystycznych i estymacji jest mniej intuitywna, ale bardziej realistyczna. Można ją nie tylko bronić, ale i stosować." Ten pogląd znakomitego znawcy rachunku prawdopodobieństwa jest sprzeczny z tezą niniejszej pracy, która orzeka paralelizm starej i nowej teorii.

Sformułujmy naszą tezę. Nic łatwiejszego w dzisiejszym stanie nauki, jak ściśle przestrzegać poprawności matematycznej i ograniczać się do wypowiadania zdań dających się uzasadnić *lege artis*. Ale postępowanie, które tu ogólnie nazwaliśmy kontrolą statystyczną, nie jest identyczne z formułowaniem zdań. Zmierza ono do pewnego celu i opiera się na pewnych hipotezach co do badanego przedmiotu. Gdy chcemy porównać ze sobą dwa sposoby postępowania, musimy zdać sobie sprawę z owego celu. Okaże się, że już w samym sformułowaniu celu tkwią owe hipotezy, od których nie jest wolna ani metoda klasyczna, ani koncepcja przedziału ufności, a nawet trudno sobie wyobrazić (jak się okaże) inną metodę, która by była wolna od hipotez. Okaże się także, że hipotezy metody klasycznej nie są ani o jotę mniej naturalne od hipotez nowej. Złudzenie, jakoby metoda przedziału ufności była wolna od hipotez, bierze się stąd, że stawia ona sobie z góry pewne sztuczne zagadnienie, które rozwiązuje poprawnie. Tą drogą dochodzi ona do pewnego postępowania, o którym można wypowiedzieć pewne prawdziwe sądy, ale te sądy nie są bynajmniej dowodem na to, że cel postępowania został osiągnięty lepiej niż przy metodzie klasycznej; żadna z metod nie ma zasadniczego prymatu.

Bez przykładów czytelnik mniej obeznany z tą problematyką tylko z trudnością połączy niniejszy wstęp z konkretnym obrazem sytuacji. Dlatego zaczniemy od starego, wprawdzie niedoskonałego, ale uproszczonego jej przedstawienia na przykładzie odbioru partii złożonej z N sztuk towaru, której przyjęcie lub odrzucenie mamy uzależnić od rezultatu zbadania próbki, czyli zbioru n sztuk wylosowanych z partii. Partia składa się z dobrych i złych sztuk; zakładamy, że badanie rozstrzyga jednoznacznie, czy zbadana sztuka jest dobra, czy zła. *Jakością partii* nazywamy frakcję a dobrych sztuk w niej zawartych; jest $a = d/N$, gdzie d oznacza liczbę dobrych sztuk w całej partii. Zagadnienie polega na wyznaczeniu tak zwanego *planu odbiorczego*; tu ograniczamy się do planów typu $m//n$: mamy pobrać próbkę n sztuk i przyjąć partię, jeżeli w próbce znajdziemy co najwyżej m złych sztuk, a odrzucić partię, jeżeli w próbce znajdziemy takich sztuk więcej niż m. Zagadnienie redukuje się do wyznaczenia liczb m i n. Widać natychmiast, że nie jest ono jeszcze dostatecznie określone, bo każda para takich liczb całkowitych m, n, że jest $0 \leqslant m \leqslant n$, daje plan $m//n$. Wprowadźmy postulat $\mathcal{R}$, który żąda, by partia, która dała w próbce m złych sztuk na n zbadanych, miała z prawdopodobieństwem 95%

jakość wyższą niż α_1, taka zaś, która dała $m+1$ złych sztuk, miała z prawdopodobieństwem 95% jakość niższą niż α_2. Gdy graniczne jakości $\alpha_1 < \alpha_2$ są z góry dane, będzie można jednoznacznie określić plan $m//n$ spełniający postulat $\mathcal{R}$, jeżeli założymy dodatkową hipotezę H, którą za chwilę wypowiemy. To jest metoda klasyczna. Wymaga ona obliczenia prawdopodobieństwa przyczyn za pomocą reguły Bayesa, a do tego potrzebna jest znajomość tak zwanych *prawdopodobieństw a priori*. Hipoteza H postuluje, że prawdopodobieństwo *a priori*, to znaczy przed zbadaniem próbki, iż jakość α partii nie przekracza β, jest β (dla każdego β z przedziału (0,1)).

Jest inna droga do sprecyzowania zagadnienia. Zamiast postulatu $\mathcal{R}$ postawimy inny, $\mathcal{P}$: Plan $m//n$ ma być taki, by partia o jakości α_2 miała prawdopodobieństwo 95%, iż zostanie przyjęta, a partia o jakości α_1 miała prawdopodobieństwo 95%, iż zostanie odrzucona. (Przedział (α_1, α_2) nie jest niczym innym jak przedziałem ufności w terminologii Spławy-Neymana.) To jest metoda nowoczesna, niemal wyłącznie stosowana w zagranicznych normach kontroli statystycznej. Nazwiemy ją *prospektywną* ($\mathcal{P}$), a poprzednią *retrospektywną* ($\mathcal{R}$).

J. Oderfeld zauważył, że przedział (α_1, α_2), który odpowiada planowi $m//n$ według interpretacji $\mathcal{P}$ tego planu, jest zbliżony do przedziału, który odpowiada temuż planowi w interpretacji $\mathcal{R}$. Znalazł on także łatwy sposób konstrukcji planów wzajemnie sobie odpowiadajacych w obu metodach, tak żeby przedział był ten sam: liczby m, n zmieniają się o jednostki[2]). Ale nas interesuje tu więcej jego trzecie spostrzeżenie: Ograniczenie się w metodzie $\mathcal{P}$ do stwierdzenia, jaką decyzję da plan w przypadku partii o jakości α_1 lub α_2, jest równoważne z pewną hipotezą *a priori*, to znaczy z pewną informacją o partii przed podjęciem badania.

Praca Oderfelda zachęciła mnie do rozważań, które ogłosiłem (w tym samym zeszycie Colloquium Mathematicum) jako krótką notę[3]). Obie noty, mojego poprzednika i moją, można jednak uważać tylko za rekonesans, a nie za wyprawę w ten teren, na który obecnie wkroczymy, a to dlatego, że nie ma w nich mowy o celu

---

[2]) J. Oderfeld, *On the dual aspect of sampling plans*, Colloquium Mathematicum II. 2 (1951), str. 89-97.

[3]) H. Steinhaus, *Quality control by sampling (A plea for Bayes' rule)*, Colloquium Mathematicum II. 2 (1951), str. 98-108.

postępowania, zwanego kontrolą statystyczną. Nie jest bowiem zadaniem kontroli statystycznej stwarzanie sposobności do wygłaszania pewnych orzeczeń o charakterze probabilistycznym. Gdyby tak było, to należało by przyznać prymat metodzie ($\mathcal{P}$), bo plany obliczane tą metodą pozwalają formułować prościej takie orzeczenia. Zawdzięczają one jednak tę zaletę zwężeniu pola widzenia do sytuacji, której sztuczność okazuje się dopiero po wytłumaczeniu empirycznego sensu orzeczeń; ta naturalna i jednolita interpretacja wymaga skierowania oka na swoisty cel postępowania.

§ 1. Cel statystycznej kontroli jakości i związane z nią straty. SKJ[4]) rozstrzyga, czy należy partię przyjąć, czy odrzucić; pomijamy tu inne możliwości, bo chodzi tylko o wyjaśnienie zasad. Tak zwany plan odbiorczy wiąże tę decyzję z rezultatami zbadania sztuk próbnych, które mogą być tylko *dobre* albo *złe* (znowu rozmyślna symplifikacja). Całe postępowanie powinno być takie, żeby wynikająca z niego *strata* gospodarcza była minimalna.

Strata składa się z dwóch dodajników: 1° szkoda wynikająca z decyzji, 2° koszt badania sztuk próbnych. Pierwszy dodajnik bierze się stąd, że badanie próbne nie daje kompletnej informacji o jakości partii. Gdybyśmy znali jakość α, moglibyśmy zawsze rozstrzygać trafnie. Jest mianowicie korzyść gospodarcza z przyjęcia partii zależna od α; ta korzyść (wyrażalna w pieniądzach) może być dodatnia lub niedodatnia, ale jest funkcją α. To samo można powiedzieć o korzyści gospodarczej z odrzucenia. Zwykle jest tak, że korzyść z przyjęcia wzrasta wraz z α, korzyść z odrzucenia maleje, gdy α rośnie. Gdybyśmy znali α, nie potrzebowalibyśmy badać sztuk próbnych, gdyż zawsze zadecydowalibyśmy o partii w sposób dający większą korzyść (w razie równości przyjmowalibyśmy też partię). To byłaby *trafna* decyzja. Porównajmy skutki takiej decyzji ze skutkami decyzji wskazanej przez plan odbiorczy. Idealny plan odbiorczy zawsze wskazywałby trafną decyzję. Z wyjątkiem pewnych skrajnych przypadków, które są bez znaczenia praktycznego, takich planów się nie spotyka. Bierze się to stąd, że plany takie wymagałyby tak wielkiej liczby sztuk próbnych, że dodajnik 2° zamieniłby w każdym przypadku całą korzyść gospodarczą na stratę. Musimy się zatem liczyć w praktyce zawsze z ewentualnością, że plan wskaże decyzję *błędną*, to jest przeciwną do trafnej. Otóż *szkodą* wynikającą

---

[4]) Ten skrót oznacza Statystyczną Kontrolę Jakości.

z decyzji (przy odbiorze według planu) czyli dodajnikiem 1° nazywamy *różnicę* (która jest zawsze nieujemna) między korzyścią z trafnej decyzji (odjemną) a korzyścią z aktualnej decyzji (odjemnikiem). Z tej definicji wynika, że gdy aktualna decyzja (czyli decyzja wskazana przez plan) jest trafna, to szkoda z decyzji jest zerem. Jak się oblicza w praktyce szkodę z decyzji, gdy aktualna decyzja jest błędna, pokażemy na przykładzie:

Zjednoczenie Przemysłu Fermentacyjnego ma własną fabrykę beczek. Fabryka posyła do browaru N beczek. Browar bada pewną liczbę z nich na szczelność; w razie odrzucenia partii wszystkie N beczek wrócą do fabryki i zostaną przesortowane sztuka w sztukę, po czym partia zostanie uzupełniona i browar dostanie N szczelnych beczek. W razie przyjęcia pierwszej partii Zjednoczenie poniesie szkodę przez to, że piwo wycieknie z nieszczelnych beczek; tej szkody nie będzie wcale w razie odrzucenia partii. W tym razie jednak Zjednoczenie poniesie koszt przesortowania całej partii, koszt dwukrotnego transportu dobrych beczek i jednokrotnego złych; tej szkody nie byłoby w razie przyjęcia partii. W obu przypadkach browar będzie mógł wyekspediować N beczek piwa; obliczenie korzyści z tego faktu dla Zjednoczenia Przemysłu Fermentacyjnego może być trudną kalkulacją, ale na szczęście tutaj zbędną; przy obliczaniu szkody z decyzji (czyli 1°) ta pozycja figuruje w odjemnej i w odjemniku i znosi się. Szkoda B z utraty piwa zależy od liczby $N(1-\alpha)$ nieszczelnych beczek, koszt C przesortowania partii zależy od N, koszt U uzupełnienia partii do stanu N jest równy wartości $N(1-\alpha)$ beczek nieszczelnych (minus wartość tej części materiału, którą można jeszcze zużyć), w końcu koszt T transportu jest $kN(1+\alpha)$, gdzie k oznacza koszt transportu jednej beczki w jedną stronę. Wszystkie wielkości B, C, U i T są znanymi funkcjami jakości α, są to funkcje liniowe, których współczynniki są znane i stałe — przyjmujemy tu też stałość liczby N. Jeżeli B jest większe od $C+U+T$, odrzucenie będzie trafną decyzją, w przeciwnym razie przyjęcie będzie trafną decyzją. Znak nierówności zależy od α i istnieje graniczne $\alpha=\bar{\alpha}$ oddzielające partie, które należy przyjmować $(\alpha \geqslant \bar{\alpha})$ od tych, które należy odrzucać $(\alpha < \bar{\alpha})$. Zjednoczenie Przemysłu Fermentacyjnego może obliczyć to $\bar{\alpha}$ raz na zawsze, ale nie może z tej wiedzy skorzystać, bo nie zna nigdy aktualnego α. Szkoda z błędnej decyzji wynosi zawsze $|B-C-U-T|$, szkoda z trafnej — *zero* (jak stwierdziliśmy wyżej). Wobec tego

szkoda z decyzji jest jedną z dwóch znanych funkcyj prawdziwej jakości a, a mianowicie $|B-C-U-T|$ lub 0. Którą z nich jest w danym przypadku, zależy od losu, albowiem od tego, jakie beczki (i w jakiej kolejności) znajdą się w próbce, będzie zależało, czy decyzja wskazana przez plan odbiorczy będzie trafna, czy błędna. Wobec tego szkodę wynikającą z decyzji (1°) oznaczymy przez $f(a,\tau)$, gdzie τ symbolizuje parametr zawisły od losu; dla pewnych wartości τ szkoda $f(a,\tau)$ jest równa $|B-C-U-T|$, dla innych — 0.

Bliższa analiza ekonomiczna prowadzi do wyrażeń bardziej skomplikowanych, co jednak nie podważa dalszych wywodów, które obejmują nie tylko uproszczony przykład.

§ 2. Koszt badania i sposób pobierania sztuk do próbki. Koszt zbadania n sztuk próbnych jest $W+cn$, gdzie W oznacza koszty wstępne, a c jest bieżącym kosztem badania jednej sztuki. Zauważmy, że przy porównywaniu planów odbiorczych W nie gra żadnej roli, bo figuruje w każdym planie (pomijamy plany, które decydują o partii bez badania) i nie wpływa na to, który plan okaże się najkorzystniejszy gospodarczo. Przypuśćmy, że w przykładzie beczek jest $N=1000$, a beczki są numerowane od 000 do 999. Niech obowiązuje plan 2//30 (por. wstęp). Z listy zawierającej numery od 000 do 999 będziemy musieli wylosować 30 numerów, napełnić wodą 30 beczek oznaczonych wylosowanymi numerami i stwierdzić po 24 godzinach, z ilu beczek woda wyciekła. Plan 2//30 każe przyjąć partię, jeżeli takich nieszczelnych beczek było 0, 1 lub 2, a odrzucić ją, jeżeli było ich 3 lub więcej. W tym przykładzie koszt badania był znany, zanim je przeprowadziliśmy — wynosi on $W+30c$ i nie zależy od a. Przypuśćmy jednak, że mamy do dyspozycji aparat złożony z pompy i manometru, tak że możemy z szybkości spadku ciśnienia powietrza wtłoczonego do beczki rozpoznać, czy beczka jest szczelna, czy nie. Będziemy badali po kolei wylosowane beczki i może się zdarzyć, że już wśród pierwszych 10 znajdziemy 3 nieszczelne. Oczywiście badanie dalszych 20 będzie wtedy zbędne i zaoszczędzimy 2/3 kosztów bieżących. Ta ewentualność mogłaby się zjawić także przy badaniu beczek wodą, gdybyśmy badali je po kolei, a nie równocześnie; oczywiście rozciągnęłoby to czas badania do tygodni i zamiast zmniejszyć zwiększyłoby koszty; przy badaniu pneumatycznym kolejność narzuca się przez to, że mamy jeden aparat, a badanie jest krótkie. Ten sposób czyni z kosztów badania

zmienną losową, bo liczba zbadanych beczek może przyjąć każdą wartość od 3 do 30; będzie to zależało od tego, na jakich numerach rozmieszczone są w partii beczki nieszczelne i jakie numery wylosujemy. Musimy wobec tego poświęcić kilka słów losowaniu.

Losowanie odbywa się w sposób gwarantujący każdej sztuce w partii jednakowe prawdopodobieństwo, że zostanie wylosowana. W naszym przykładzie $N=1000$, więc to prawdopodobieństwo jest 1/1000. Rzucamy trzy kostki graniaste dziesięciościenne; każda ma na ścianach cyfry $0,1,2,\ldots,9$, po jednej na każdej ścianie, a na podstawach kostek są cyfry I, II, III odróżniające kostki od siebie. Kostka I da pierwszą, kostka II drugą, a kostka III trzecią cyfrę. Tym sposobem każdy rzut da nam jeden z numerów od 000 do 999. Zwykła teoria planów pojedynczych $m//n$ przewiduje, że przy wyjściu powtórnym tego samego numeru zapisuje się rezultat badania wylosowanej sztuki, tak jak gdyby to była nowa sztuka, choć bada się ją tylko raz; ta zasada nie jest wcale konieczna, ale upraszcza ona teorię matematyczną, a na nasze rozważania nie ma żadnego wpływu. Sprawia ona jednak, że koszt badania będzie zawsze zmienną losową; przy badaniu beczek sposobem hydraulicznym i przy planie 2//30 może się też zdarzyć, że naprawdę zbadamy tylko 10 beczek. Tak będzie, gdy losowanie da przy 30 rzutach kostkami tylko 10 różnych numerów, bo wtedy tylko 10 beczek napełnimy wodą. (Ta paradoksalna możliwość świadczy o tym, jak daleko odeszła SKJ od intuicji laika.)

Oczywiście, można by określić plany pojedyncze inaczej i rozumieć symbol $m//n$ w ten sposób, że pomija się rzuty kostką, które powtarzają numery wyszłe poprzednio — w przykładzie 2//30 rzucałoby się tak długo, żeby wyszło 30 różnych numerów, a nie 30 razy. Teoria matematyczna takich planów byłaby inna, ale i w tym przypadku można uznać koszt badania za zmienną losową korzystając z tego, że pojęcie zmiennej obejmuje stałą. Przemawia za takim ujęciem kosztu badania, że przy niektórych badaniach złe sztuki ulegają zniszczeniu, a dobre nie; jeżeli złe sztuki nie są zupełnie bezwartościowe, koszt badania staje się zależny od ilości takich sztuk w próbce.

Zwróćmy uwagę na to, że przy ustaleniu w umowie o dostawę i odbiór nie tylko planu odbiorczego, ale i sposobu losowania, np. losowania kostkami, staje się obojętne rozmieszczenie złych sztuk w partii w tym sensie, że nie może ono ani zwiększyć, ani zmniejszyć szans dostania się do próbki którejkolwiek określonej złej

sztuki. Wobec tego przyjmiemy raz na zawsze, że wszystkie sztuki o numerach $0,1,2,\ldots,(d-1)$ są dobre ($d=Na$). Przy tym założeniu (które nie uszczupla ogólności) jakość a i kolejność wylosowanych numerów determinuje kompletnie zarówno koszta badania jak decyzję, którą orzecze ustalony w umowie plan P.

Chcielibyśmy tu nawiasowo wspomnieć o tym, że ulepszenie zwane *analizą sekwencyjną*, które pozwala na ewentualne skrócenie badania (czyli zmniejszenie liczby sztuk próbnych), a które wyjaśniliśmy pobieżnie na przykładzie beczek, daje się bez żadnych trudności wprowadzić do metody klasycznej. Zamiast tak zwanego planu pojedynczego $m//n$ wprowadzamy przepis, który każe badać sztuki kolejno (w porządku wylosowanych numerów) i dla każdej liczby naturalnej n zbadanych sztuk obliczać z niej i z liczby m złych sztuk wśród nich znalezionych prawdopodobieństwo $p_1(m,n)$, że partia ma jakość wyższą niż a_2 i prawdopodobieństwo $p_2(m,n)$, że ma jakość niższą niż a_1; badanie kończy się na takim n, dla którego któraś z liczb p_1,p_2 przekroczy 0,95. Jeżeli jest nią p_1, partię się przyjmuje, jeżeli p_2 — odrzuca. Oczywiście hipoteza H metody ($\mathcal{R}$) (por. wstęp) zostaje utrzymana. Tę uwagę umieszczamy tu dlatego, że niektórzy statystycy nazywają — za Abrahamem Waldem — analizą sekwencyjną postępowanie, które łączy metodę prospektywną ($\mathcal{P}$) z postulatem najkrótszego badania. To budzi błędne mniemanie, jakoby zaleta najkrótszego (a więc najtańszego) badania była istotnie związana z nową metodą i była dowodem jej wyższości nad klasyczną. Jest wprost przeciwnie: hipoteza H pozwala sformułować w kilku słowach, jak to uczyniliśmy o parę wierszy wyżej, metodę retrospektywno-sekwencyjną; metoda prospektywno-sekwencyjna Walda wymagała dużej pomysłowości, a jej sformułowanie i wyprowadzenie po dziś dzień nie jest kompletne. Dodajmy tutaj, że Komisja SKJ Polskiego Komitetu Normalizacyjnego rozpoczynając swoją działalność w 1947 r. sformułowała metodę retrospektywno-sekwencyjną jako naturalnie się narzucającą.

§ 3. Dalsze sformalizowanie zagadnienia SKJ. Powtórzmy konkluzje z poprzednich paragrafów. Strata związana z SKJ składa się z dwóch dodajników: ze szkody wynikającej z decyzji i z kosztu badania sztuk próbnych. Pierwszy dodajnik ma postać $f(a,\tau)$, gdzie parametr τ symbolizuje zależność od losu, a postać funkcji f jest znana, drugi, który nazwiemy $g(a,\tau)$, jest także znaną funkcją tych dwóch zmiennych. Wobec tego

całkowita strata, jako suma tych dwóch funkcyj, także ma postać funkcji zmiennych α i τ. Zauważmy teraz, że cały przebieg SKJ jest kompletnie zdeterminowany, gdy (przy ustalonym planie odbiorczym P) dana jest jakość partii α i ciąg wylosowanych numerów, gdyż wszystkie złe sztuki mają numery od d włącznie do N wyłącznie. Wobec tego parametr losowy τ zależy tylko od α i od ciągu wylosowanych numerów: $\tau = \varphi(\alpha, t)$, gdzie t symbolizuje ów ciąg. Wstawiając tę funkcję φ za τ w sumę $f+g$, wyrazimy całkowitą stratę przez znaną (przynajmniej teoretycznie) funkcję $S(\alpha, t)$. Tutaj t symbolizuje ciąg wylosowanych numerów, ale można będzie t uznać za liczbę, jeżeli uda się nam przypisać każdemu ciągowi numerów reprezentanta liczbowego. To łatwe zadanie tak się rozwiązuje dla $N = 1000$ (i analogicznie ogólnie): Bez względu na to, ile losowań wymaga plan, możemy wyobrazić sobie, że maszyna losująca przy każdym puszczeniu jej w ruch produkuje nieskończony ciąg $\{v_n\}$ trójcyfrowych numerów. Temu ciągowi przypiszemy liczbę t daną przez rozwinięcie $0, v_1 v_2 v_3 \ldots v_k \ldots$ odczytane w układzie tysiączkowym. Chcąc, by tak określona odpowiedniość między liczbami a ciągami była jednoznaczna, musimy założyć, że maszyna nigdy nie da ciągu, który by począwszy od pewnego miejsca składał się z samych numerów 000. Ponieważ prawdopodobieństwo wyprodukowania takiego ciągu jest zero, więc to założenie w niczym nie wpłynie na dalsze wywody. Wtedy każdej liczbie z przedziału $(0,1)$ będzie odpowiadał ciąg, a każdemu ciągowi — liczba. Przy tym odpowiedniość ta ma następującą ważną właściwość:

Prawdopodobieństwo, że ciąg numerów ma dowolną własność W, jest równe mierze (Lebesgue'a-Borela) zbioru tych t, które są przypisane ciągom numerów o własności W.

Dowód wygłoszonego tu twierdzenia opiera się na założeniu jednakowego prawdopodobieństwa wyjścia każdego z numerów od 000 do 999; w ogólnym przypadku (N dowolne) jest analogiczny.

Dla wyjaśnienia związku między ciągiem $\{v_n\}$ a decyzją wprowadzamy ciągi

$$x_1, x_2, \ldots, x_n, \ldots \tag{1}$$

$$r_1, r_2, \ldots, r_n, \ldots, \tag{2}$$

gdzie x_n jest 1, gdy sztuka oznaczona numerem v_n jest dobra,

a jest 0, gdy ta sztuka jest zła; r_n oznacza liczbę różnych numerów wśród numerów $v_1, v_2, \ldots, v_n$ (ciąg $\{r_n\}$ jest niemalejący).

Plan odbiorczy będzie zdefiniowany, gdy określimy ciąg funkcyj

(3) $d_0, d_1(x_1; r_1), d_2(x_1, x_2; r_1, r_2), \ldots, d_n(x_1, x_2, \ldots, x_n; r_1, r_2, \ldots, r_n), \ldots,$

tak zwanych *funkcyj decyzyjnych*, dla wszystkich ciągów (1) i (2). Wartości tych funkcyj są nakazami $\mathcal{A}$, $\mathcal{B}$ lub $\mathcal{C}$, przy czym

(4)
$\mathcal{A} \equiv$ przyjąć partię,
$\mathcal{B} \equiv$ odrzucić partię,
$\mathcal{C} \equiv$ zbadać sztukę o numerze v_{n+1} i wykonać nakaz d_{n+1} $(n = 0, 1, 2, \ldots)$.

Funkcje d_n mają spełniać następujące relacje:

(5.1) jeżeli dla jakiejś pary ciągów (1) i (2) jest $d_n \not\equiv \mathcal{C}$, to dla tejże pary jest $d_{n+1} \equiv d_n$;

(5.2) jeżeli dla jakiejś pary ciągów (1) i (2) jest $d_n \equiv \mathcal{C}$, a $r_{n+1} = N$, to dla tejże pary jest $d_{n+1} \not\equiv \mathcal{C}$ $(n = 0, 1, 2 \ldots)$.

Ponieważ każde t $(0 < t \leqslant 1)$ określa jednoznacznie $\{v_n\}$, a zatem (1) i (2), więc można funkcje decyzyjne d_n oznaczać przez $d_n(t)$.

Słowna interpretacja postulatu (5.1). Ciąg nakazów uzyskanych z (3) przez wstawienie rezultatów losowania $\{r_n\}$ i rezultatów badania $\{x_n\}$ jest obliczalny i zgodny: dopóki pojawia się odraczająca decyzja $\mathcal{C}$, dopóty kontynuujemy badanie i uzyskujemy wartości x i r potrzebne do powzięcia następnej decyzji, a gdy pojawi się jedna z decyzyj kategorycznych $\mathcal{A}$ lub $\mathcal{B}$, to zbędne jest dalsze badanie, bo wartości wszystkich następnych d są identyczne z tą decyzją, jakiekolwiek byłyby dalsze, nieznane x i r.

Słowna interpretacja postulatu (5.2). Gdy liczba różnych sztuk zbadanych zrówna się z N (to znaczy, gdy poznamy prawdziwą jakość partii), to zjawia się decyzja kategoryczna, o ile nie zjawiła się wcześniej.

Niesprzeczność postulatów wynika z trywialnych przykładów. Gdy np. wszystkie d_n $(n = 0, 1, 2, \ldots)$ mają wartość $\mathcal{A}$, postulaty są oczywiście spełnione. Warto jednak podać przykład nietrywialny, a mianowicie ująć w schemat (3) sekwencyjny wariant planu 2//30 z § 2. W tym celu przyjmijmy $r_0 = 0$ i zauważmy, że wśród sztuk o numerach $v_1, v_2, \ldots, v_n$ jest r_n różnych, a z nich jest

$$a_n = \sum_{k=1}^{n} (r_k - r_{k-1}) x_k$$

dobrych i

$$z_n = \sum_{k=1}^{n} (r_k - r_{k-1})(1 - x_k)$$

złych.

Nazwijmy ν najmniejsze n, dla którego zachodzi jedna z relacyj $a_n = 28$, $z_n = 3$. Łatwo spostrzec, że dla $n = \nu$ zachodzi *tylko* jedna z nich; jest oczywiście $\nu \geqslant 3$ (jeżeli ν istnieje). Określamy: $d_n \equiv \mathcal{C}$ dla $n < \nu$, $d_\nu \equiv \mathcal{A}$, gdy zachodzi pierwsza relacja, $d_\nu \equiv \mathcal{B}$, gdy druga i $d_n \equiv d_\nu$ dla $n \geqslant \nu$. Czytelnik łatwo sprawdzi, że tak określony plan typu (3) spełnia (4) i (5), a ponadto odpowiada sekwencyjnemu wariantowi planu 2//30 i to takiemu, przy którym sztuki powtórnie wylosowane nie mają wpływu na decyzję.

Zwykle definiuje się plany odbiorcze prościej, bez uwzględniania ciągu $\{r_n\}$; poprzestaje się na funkcjach decyzyjnych zależnych od $\{x_n\}$. Ponieważ zależność od obu ciągów obejmuje jako przypadek szczególny zależność od jednego, więc sama forma definicji (3) nie jest uszczupleniem ogólności. Jest nim jednak postulat (5.2), który uzależnia plan od liczności partii N. Słowna interpretacja tego postulatu wyjaśnia, że jest to przepis, który znajduje zastosowanie tylko wtedy, gdy się okaże, że zbadaliśmy całą partię. Wtedy jednak najlepszą decyzją, jaką można powziąć, jest przyjęcie partii, gdy prawdziwa jakość (którą poznaliśmy) jest niemniejsza niż $\bar{a}$ określone w § 1, odrzucenie, gdy ta jakość jest mniejsza niż $\bar{a}$. Ponieważ $\bar{a}$ jest znane *a priori*, można tak zdefiniować funkcje d_n w (3), żeby w owym wyjątkowym przypadku decyzja była najlepsza. Ten postulat nie ogranicza zupełnie swobody definiowania decyzyj w innych przypadkach. Wobec tego można ulepszyć każdy plan typu (3) spełniający (4) i (5.1) przez dodanie (5.2). Rozumiemy to w ten sposób, że dołączenie (5.2) nie zmienia decyzji w żadnym przypadku, z wyjątkiem jednego, w którym wyłącza możliwość błędnej decyzji nie przedłużając badania. Postulat (5.2) nie wchodzi bowiem w grę, jeżeli była decyzja kategoryczna d_n, a jeżeli jej nie było, to i bez tego postulatu nastąpiłoby badanie sztuki nr v_{n+1}, którym zadowala się postulat. Postulat nie może więc nigdy zwiększyć całkowitej straty gospodarczej. Ponieważ jednak naszym celem jest wyszukanie planów optymalnych pod względem

gospodarczym, więc zastąpienie każdego planu P przez inny, co najmniej taki dobry jak P, nie oddali nas od tego celu.

Zauważmy teraz, że postulat (5.2) gwarantuje decyzję kategoryczną w każdym przypadku, w którym w ciągu $\{r_n\}$ pojawi się N. Jakie jest prawdopodobieństwo, że to się zdarzy? Prawdopodobieństwo, że określony numer v' $(0 \leqslant v' < N)$ nie pojawi się w skończonej serii $v_1, v_2, \ldots, v_n$ ani razu, jest $(1 - 1/N)^n$; że wogóle jakiś z numerów $0, 1, \ldots, N-1$ się w niej nie pojawi, jest niewiększe niż $N(1 - 1/N)^n$. Stąd prawdopodobieństwo, że w nieskończonym ciągu $\{r_n\}$ zabraknie N, będzie niewiększe niż

$$\lim_{n \to \infty} N\left(1 - \frac{1}{N}\right)^n,$$

który to *limes* jest 0. Wobec tego zbiór liczb t reprezentujących ciągi $\{v_n\}$ nie wyczerpujące wszystkich numerów w partii, ma miarę zerową. Innymi słowy: dla *prawie wszystkich* t ciągi numerów symbolizowane przez t prowadzą — przy ustalonym N i ustalonym planie badania P typu (3) (4) (5) — do kategorycznej decyzji.

Ponieważ koszt badania nie przekroczy nigdy kosztu badania N sztuk, a więc pewnej liczby skończonej, a także szkoda wynikająca z decyzji jest ograniczona, więc całkowita strata jest ograniczoną, nieujemną funkcją zmiennych a i t. Ta funkcja jest ciągła względem t na zbiorze Z tych t, które dostarczają kategorycznych decyzyj (a więc na zbiorze o mierze 1). Dowód polega na tym, że przy ustalonym a oba dodajniki (szkoda z decyzji i koszt badania) zależą tylko od t; jeżeli $t_1 \epsilon Z$, a ponadto ma tylko jedno rozwinięcie, to istnieje takie n, że $d_n(t_1) \neq \mathcal{C}$. Napiszmy istotnie nieskończone rozwinięcie $t_1 = 0, v_1 v_2 \ldots v_n \ldots$ i przypuśćmy, że t' leży tak blisko t_1, że jego istotnie nieskończone rozwinięcie nie różni się aż do n-tego znaku włącznie od rozwinięcia t_1 (tak musi być przy dostatecznej bliskości). Wtedy będzie $d_k(t') \equiv d_k(t_1)$ dla $k \leqslant n$, a wobec $d_n(t_1) \neq \mathcal{C}$ i (5.1) także dla $k > n$; będzie $t' \epsilon Z$, a ponadto zarówno decyzja jak liczba sztuk zbadanych będzie ta sama dla t' co dla t_1; całkowita strata będzie zatem, jako funkcja t, stała w pewnym przedziale otaczającym t_1, a tym bardziej ciągła w t_1, c. b. d. o.

Okazuje się, że całkowita strata $S(a,t)$ jest przy ustalonym a funkcją zmiennej t określoną w pewnym zbiorze o mierze 1, ograniczoną, nieujemną i ciągłą w każdym punkcie tego zbioru, a zatem całkowalną według Riemanna względem t $(0 < t \leqslant 1)$.

Jakość a może przyjmować tylko wartości $0, 1/N, 2/N, \ldots, N/N$. Dla ułatwienia wprowadźmy konwencję (nie odbiegającą od rzeczywistości), że a nie może być zerem. Określmy $S(a,t)$ dla wszystkich a od zera (wyłącznie) do 1 (włącznie): dzieląc przedział $(0,1\rangle$ na N równych przedziałów I_k $(k=1,2,\ldots,N)$, zaliczmy do każdego I_k jego prawy koniec β_k i przyjmijmy $S(a,t)=S(\beta_k,t)$, gdy $a \in I_k$. Określona w ten sposób funkcja $S(a,t)$ jest (dla każdego $t \in Z$) ciągła względem a z wyjątkiem $N-1$ punktów. Ilekroć będziemy mówili, że partia ma jakość a, będziemy to rozumieli nie dosłownie, lecz w tym sensie, że frakcja d/N dobrych sztuk w partii spełnia nierówność $d/N \geqslant a > (d-1)/N$. Dzięki tej umowie i rozszerzeniu definicji funkcji S uchronimy się od nieścisłości w dalszych wywodach, gdzie będzie trzeba uważać a za zmienną ciągłą.

§ 4. Postulat minimalnej straty całkowitej. Jak zaznaczyliśmy już w § 1, celem SKJ jest zminimizowanie straty, która powstaje zarówno przy odbiorze partii zbadanej sztuka w sztukę, jak przy odbiorze bez badania. W pierwszym przypadku strata wynika z kosztów badania, w drugim z decyzji o partii niewiadomej jakości. Pomysł SKJ jest naturalny: szukanie optimum w postępowaniu pośrednim, to jest w zbadaniu próbki i uzależnieniu decyzji od wyniku badania. W poprzednich paragrafach określiliśmy i scharakteryzowaliśmy matematycznie funkcję S, która wyraża całkowitą stratę. Nie interesowaliśmy się jednak jej zależnością od planu odbiorczego P. Teraz musimy stratę zapisać w postaci $S(P;a,t)$; S jest funkcjonałem planu P, a dla ustalonego P zamienia się w znaną funkcję dwóch parametrów: jakości a i parametru losowego t. Postulat zminimizowania S przez dobór optymalnego planu P nabiera określonego sensu dopiero po pozbyciu się tych parametrów. Najprościej byłoby zażądać, żeby P minimizowało S dla każdego a i t. Są specjalne przypadki, w których to jest możliwe: partia towaru niezbędnego, z której odrzucenia wynikłaby szkoda większa od kosztu przesortowania całej partii, a której jakość ma małe znaczenie dla odbiorcy; tutaj plan P_0 określony przez $d_n \equiv \mathcal{A}$ dla wszystkich n (według (3) ma żądaną własność, bo jest $S(P_0,a,t) < S(P,a,t)$ dla wszystkich a i t przy $P \neq P_0$. Wyjątkowość tego przykładu jest wyraźna.

Usunięcie parametru t przez rozpatrywanie oczekiwanej straty $E(S)$ zamiast prawdziwej, jest powszechnie przyjęte. Dezyderat minimalnej straty przyjmie postać

$$(6) \qquad \int_0^1 S(P,a,t)\,dt = \text{minimum} \quad \text{dla} \quad P = P_0,$$

z którego to warunku należy wyznaczyć P_0. Jaki jest realny sens tego warunku? Wyobraźmy sobie, że odbieramy bezustannie tę samą partię według tego samego planu P. Za każdym razem aparat losujący dostarczy nam jakiejś liczby $t = 0,v_1v_2\ldots v_k\ldots$ (por. § 3), a więc ciąg odbiorów wytworzy ciąg $\{t_n\}$ takich liczb. Właściwość zmiennej losowej t sformułowana w § 3 określa jej rozkład: jest on równomierny, to znaczy, że prawdopodobieństwo, iż wylosujemy t należące do odcinka $\langle 0,T\rangle$, jest T (dla każdego T między 0 a 1). Każde t_n ma taki rozkład równomierny, a ponieważ losowanie przy danym odbiorze jest niezależne od losowań przy innych odbiorach, więc t_n są wzajemnie niezależne. Z tych własności ciągu zmiennych losowych $\{t_n\}$ wynika, że z prawdopodobieństwem 1

$$(7) \qquad \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} S(P,a,t_j) = \int_0^1 S(P,a,t)\,dt.$$

Praktycznie można uważać za pewne to, co ma prawdopodobieństwo 1, więc eliminacja całki z (6) i (7) zamieni dezyderat minimalnej straty oczekiwanej na żądanie, żeby średnia strata przy nieskończenie wielu odbiorach tej samej partii według planu P_0 była co najwyżej taka, jak owa średnia przy odbiorach według jakiegokolwiek planu P. To sformułowanie (oparte na znanym wariancie (7) prawa wielkich liczb) usprawiedliwia dezyderat (6): wyrównanie gry przypadku przez wzięcie średniej straty przy pomyślanym powtarzaniu się tej samej sytuacji jest naturalnym zabiegiem i jest przyjęte nie tylko w SKJ. Udaje się on tutaj dzięki temu, że używamy maszyny do losowania, która wybiera każdy numer v z prawdopodobieństwem $1/N$, co gwarantuje bez dodatkowych hipotez równomierny rozkład parametru t. (Ta równomierność uzasadnia *ex post* nazwanie całki w (6) oczekiwaną stratą $E(S)$, jednak pojęcie *oczekiwanej* straty ma tu tylko werbalne znaczenie.) Całkę (6) należy rozumieć w sensie Riemanna. Usunięcie parametru t przez całkowanie nie daje jeszcze określonego sensu żądaniu minimalnej straty, bo pod całką (6) figuruje jeszcze a. Zadanie (6) ma na ogół przy każdym a inne rozwiązanie P_0, a ponieważ w praktyce a jest nieznane, więc wyszukanie P_0 nie jest możliwe bez eliminacji a. (Znowu pomijamy ekscentryczne przykłady niezależności straty od a.)

§ 5. Eliminacja jakości a. Przejdziemy po kolei kilka sposobów.

I. Usuwamy a przez całkowanie. Dezyderat (6) przyjmuje postać

$$(8) \qquad \int_0^1\int_0^1 S(P,a,t)\,da\,dt = \text{minimum} \qquad \text{dla} \quad P = P_0,$$

z którego to warunku należy wyznaczyć P_0. To zadanie ma zupełnie określony charakter i wyznacza P_0 we wszystkich niemal przypadkach spotykanych w praktyce, jakkolwiek trudności techniczne efektywnego znalezienia tego optymalnego planu mogą być nieraz niemałe. Jaki jest jednak tym razem realny sens warunku minimalnej straty? Wyobraźmy sobie znowu, że odbieramy według planu P coraz to nowe partie tego samego towaru, ale o różnych jakościach a_j. Załóżmy — hipoteza H_1 —, że ciąg $\{a_j\}$ ma *ekwipartycję*, to znaczy, że frekwencja względna wyrazów tego ciągu padających w przedział $(0,\beta)$ jest β (dla każdego β między 0 a 1). Będzie wtedy z prawdopodobieństwem 1 zachodziła relacja

$$(9) \qquad \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} S(P,a_j,t_j) = \int_0^1\int_0^1 S(P,a,t)\,da\,dt,$$

gdzie, tak jak w § 4, ciąg $\{t_n\}$ jest wytworzony przez aparat losujący.

Dowód niniejszego twierdzenia polega na tym, że wobec równomierności zmiennej losowej t ciąg $\{t_n\}$ ma ekwipartycję z prawdopodobieństwem 1, i że ciągi $\{a_n\}$, $\{t_n\}$ są od siebie niezależne (co wynika z użycia tego samego aparatu do losowania przy każdym odbiorze). Wobec tego punkty (a_n,t_n) mają z prawdopodobieństwem 1 ekwipartycję płaską w kwadracie $(0 \leqslant a \leqslant 1)$ $(0 \leqslant t \leqslant 1)$, to znaczy, że frekwencja tych spośród nich, które padają w prostokąt $(0 \leqslant a \leqslant \beta)$ $(0 \leqslant t \leqslant T)$ jest βT (dla każdego punktu (β,T) kwadratu). Ostatni krok dowodu wymaga podwójnej całkowalności riemannowskiej funkcji S, która to własność wynika z tego, że możemy kwadrat podzielić liniami $a = \beta_k$ $(k = 1,2,\ldots,N)$ na N prostokątów o podstawach I_k (por. koniec § 3), a w każdym z nich S jest stałe względem a, jedyne zaś punkty nieciągłości są to punkty nieciągłości względem t, które razem dają zbiór o płaskiej mierze 0. Zaznaczmy, że użyty tu wariant prawa wielkich liczb nie był dotychczas stosowany w te-

orii SKJ. Zwróćmy także uwagę na to, że wzory (6) i (7) i towarzyszące im wyobrażone procesy są zbędne, gdyż można było od razu usunąć obydwa parametry, jak zrobiono w (8), jednak dydaktycznie byłoby to chybione.

Eliminacja całki z (8) i (9) zamienia obecnie dezyderat minimalnej straty na żądanie, żeby średnia strata, przy nieskończenie wielu kolejnych odbiorach według planu P_0 partyj o różnych jakościach, była co najwyżej taka, jak przy odbiorach tychże partyj według jakiegokolwiek planu P, przy czym zakładamy ekwipartycję owych kolejnych jakości. Nawiasowo zaznaczmy, że hipoteza ekwipartycji H_1 wobec umowy wprowadzonej na końcu § 3 żąda, żeby frekwencja każdej z N możliwych jakości d/N $(d=1,2,\ldots,N)$ była $1/N$. Zauważmy także, że pojęcie oczekiwanej straty w sensie § 4 nie figuruje w sposobie I obecnie omawianym.

II. Usuwamy a przez maksymizację lewej strony (6) względem a, przez co dezyderat minimalnej straty przyjmuje postać żądania

$$\max_{0\leqslant a\leqslant 1} \int_0^1 S(P,a,t)\,dt = \text{minimum} \quad \text{dla} \quad P=P_0, \tag{10}$$

z którego należy wyznaczyć P_0. Zauważmy, że (10) żąda wyznaczenia naprzód takiego a, które przy danym P maksymizuje oczekiwaną stratę (czyli całkę $\int_0^1 S\,dt$); to $a=F(P)$ jest funkcjonałem planu P i zamienia oczekiwaną stratę $E(S)$ na funkcjonał planu P, a mianowicie na

$$\int_0^1 S(P,F(P),t)\,dt, \tag{11}$$

który to funkcjonał należy zminimizować przez dobór P; P_0 jest nazwą na P minimizujące (11).

Załóżmy — hipoteza H_2 —, że między planem przyjętym P a jakością a dostarczanych partyj zachodzi związek $a=F(P)$ (wyżej określony). Zastępując całkę w (10) przez lewą stronę wzoru (7) będziemy mogli dezyderat minimalnej straty zamienić na żądanie, żeby średnia strata przy nieskończenie wielu odbiorach partyj o tej samej jakości według planu P_0 była co najwyżej taka, jak średnia przy odbiorach według jakiegokolwiek planu P. (P_0 spełnia ten dezyderat z prawdopodobieństwem 1, bo takie jest prawdopodobieństwo relacji (7).) To jest — tym razem — realny sens warunku minimalnej straty.

III. Usuwamy a przez ustalenie. Niech np. $a = a_0$. Dezyderat (6) zamieni się na warunek

$$(12) \qquad \int_0^1 S(P, a_0, t)\,dt = \text{minimum} \qquad \text{dla} \quad P = P_0,$$

służący do wyznaczenia P. Niech hipoteza H_3 orzeka, że wszystkie partie mają jakość a_0. Przy tej hipotezie plan P_0 wyznaczony z (12) realizuje najmniejszą średnią stratę przy kolejnych odbiorach.

IV. Usuwamy a przez ustalenie i sumowanie skończone. Przy dwóch dodajnikach dezyderat (6) przyjmie postać

$$(13) \qquad \int_0^1 S(P, a_1, t)\,dt + \int_0^1 S(P, a_2, t)\,dt = \text{minimum} \qquad \text{dla} \quad P = P_0.$$

Niech hipoteza H_4 orzeka, że każda partia ma jakość a_1 lub a_2, przy czym każda z tych jakości pojawia się w ciągu kolejno przedstawianych partyj z frekwencją 1/2. Przy hipotezie H_4 plan P_0 wyznaczony z (13) realizuje najmniejszą średnią stratę przy kolejnych odbiorach.

V. Usuwamy a tak jak w sposobie I, jednak pod całką zamiast S piszemy S^2:

$$(14) \qquad \int_0^1\int_0^1 S^2(P, a, t)\,dt\,da = \text{minimum} \qquad \text{dla} \quad P = P_0.$$

Plan wyznaczony przez (14) zrealizuje minimum średniego *kwadratu* straty przy kolejnych odbiorach, gdy przyjmiemy hipotezę $H_5 = H_1$.

§ 6. Porównanie sposobów eliminacji parametrów i hipotez H_1-H_5. Każdy z dezyderatów § 5, a więc (8), (10), (12), (13) i (14) określa pewien plan odbiorczy. Matematyczna poprawność każdej z pięciu definicji planu P_0 nie wymaga żadnych hipotez. Każda definicja określa inny plan; nazwijmy te plany P_1, P_2, P_3, P_4 i P_5. Na razie ograniczmy się do pierwszych czterech (P_1-P_4). Ponieważ — jak podkreśliliśmy we wstępie — zagadnienie SKJ jest zagadnieniem optymalnego postępowania, sama poprawność definicji planu nie rozwiązuje go bez określenia wartości postępowania, które ów plan przepisuje. Wartość postępowania oceniamy według straty gospodarczej, jak to wytłumaczono w § 1. Wyraża się ona przez $S(P, a, t)$. Chcąc pozbyć się parametrów t i a musimy rozpatrywać średnią stratę przy ciągu od-

biorów. Taka średnia istnieje i przy planie P_j ($j = 1,2,3,4$) spełnia warunek minimalności, jeżeli przyjmiemy hipotezę H_j. Tak więc nieuniknione jest przyjęcie właściwej hipotezy, jeżeli chcemy nadać postulatowi optymalnego postępowania treść podległą kontroli doświadczalnej. Wobec tego porównanie sposobów I-IV redukuje się do zbadania, które z hipotez H_1-H_4 lepiej, a które gorzej odpowiadają rzeczywistości.

Hipoteza H_1 jest matematycznie prosta, ale to nie jest dostateczną legitymacją. Nic nie uprawnia do przypuszczenia, że kolejno dostarczane partie mają ekwipartycję jakości, natomiast są poważne powody do przypuszczeń, że jest inaczej. Zauważmy, że H_1 wynika z prawdopodobieństwem 1 (czyli prawie na pewno) z hipotezy H metody klasycznej (czyli $\mathcal{R}$) sformułowanej we wstępie. Hipoteza H_1 jest więc arbitralna, ale nie jest w wyraźnej sprzeczności ze zdrowym rozsądkiem.

Hipoteza H_2 zakłada, że jakość dostarczanych partyj zależy od planu P. Wprawdzie plan odbiorczy jest znany dostawcom i możnaby przypuścić, że wpływa on na jakość partyj, ale w hipotezie H_2 zależność nie jest taka, któraby zapewniała największą korzyść dostawcy: jest to zależność, w której jakość jest dobrana do planu tak, by całkowita strata gospodarcza była maksymalna. Trudno wyobrazić sobie sytuację, w której czynniki wpływające na jakość partii — mówię o czynnikach personalnych — mogą i chcą dobierać ją tak, by maksymizować stratę gospodarczą wynikającą z procesu SKJ; nawet najlepszych sabotażystów nie można o to posądzić, bo oni obniżą jakość do możliwie najniższego poziomu, ale to wcale nie da dużej *straty z SKJ*. Przeciwnie, plany obliczone sposobem II odrzucą liche partie po niedużej liczbie prób, więc dadzą trafną decyzję tanim kosztem, a tym samym nieznaczną stratę gospodarczą *przy SKJ*[5]). Hipoteza H_2 okazuje się fantastyczna na tle rzeczywistości. To osłabia znaczenie sposobu II, znanego i używanego pod nazwą „minimax".

Hipoteza H_3 ma za sobą tylko ten argument, że warunki produkcji zmuszają dostawcę do ustalenia jakości. Zauważmy, że sposób III pozwala natychmiast efektywnie wyznaczyć plan P_0 z gra-

---

[5]) Strata gospodarcza z obniżenia jakości będzie olbrzymia, ale *nie* będzie to strata powstała *z okazji SKJ*. Np. jeśli wiadomo, że wszystkie partie mają jakość 0, najlepszym planem będzie odrzucenie każdej partii bez badania; to kompletne położenie produkcji nie ucieszy naszego fikcyjnego maniaka, bo *strata z SKJ będzie zero*.

nicznego $\bar{a}$ określonego w § 1. Jeżeli $a_0 \geqslant \bar{a}$, plan P_0 każe przyjąć każdą partię bez badania, jeżeli $a_0 < \bar{a}$, plan P_0 każe odrzucić każdą partię bez badania. Rzeczywiście, warunek (12) będzie wtedy spełniony, bo decyzja będzie zawsze trafna i szkoda z decyzji będzie zerem, a koszt próbki będzie minimalny, bo nie zbada się ani jednej sztuki; S nie będzie zależało od t i całka (12) przyjmie najmniejszą wartość, którą może przyjąć przy dopuszczeniu do konkurencji wszelkich P i a. Wprowadzenie hipotezy H_3, która uzasadnia użycie sposobu III, niweczy zatem w praktyce SKJ, bo prowadzi do planu, który traci sens, gdy hipoteza nie jest spełniona, czego nie można zarzucić innym hipotezom tu omawianym. H_3 jest najgorszą z pięciu hipotez.

Hipoteza H_4 jest ogólniejsza niż H_3 — zawiera H_3 w przypadku szczególnym $a_1 = a_2$. H_4 nie podlega zarzutowi podniesionemu przeciw H_3, gdy jest $a_1 < a_0$ i $a_2 > a_0$. Dlatego też sposób IV daje plan P_0, który nie staje się absurdalny, gdy hipoteza H_4 nie odpowiada rzeczywistości. Ale ta zaleta (którą mają i inne sposoby, z wyjątkiem III) nie usprawiedliwia sztuczności hipotezy H_4, trudno bowiem podać przykład wzięty z praktyki, w którym by wszystkie partie miały zawsze jedną z dwóch znanych jakości, i to równie często jedną jak drugą. Tu już argument wysunięty w obronie hipotezy H_3 staje się kontrargumentem.

Przegląd hipotez H_1-H_4 przekonywa, że żadna z nich nie ma za sobą poparcia rzeczywistości, wszystkie bowiem są mniej lub więcej sztuczne i hipoteza H_1, w której łatwo się dopatrzeć potępionego postulatu Bayesa nie tylko nie jest gorsza od innych, lecz na ich tle przedstawia się zupełnie nieźle.

Nie zapominajmy o tym, że użyteczność sposobów I-V nie jest związana ściśle z prawdziwością hipotez. Każdy sposób definiuje pewien plan P_0, którego można używać nawet, gdy przynależna hipoteza nie jest spełniona. Hipoteza jest potrzebna tylko do dowodu, że plan ten jest optymalny (w znaczeniu minimalnej szkody z SKJ). Z jednym wyjątkiem sposobu III wszystkie nadają się do użytku praktycznego i obliczone z nich plany nie odbiegają istotnie od pospolicie używanych, chociaż praktycy SKJ niewiele troszczą się o zgodność hipotez z rzeczywistością.

Dotychczas paralelizm wszystkich sposobów jest zachowany. Każdy z nich prowadzi do optymalnego planu przy właściwej hipotezie, jeżeli dobroć planu oceniamy według średniej straty całkowitej. Wyjątek stanowi tu sposób V; przy hipotezie H_1 daje on

plan optymalny w sensie minimum średniej *kwadratu* straty. Podaliśmy sposób V jako przykład bogactwa możliwości, gdy chodzi o zamianę a na zmienną pozorną. Sposobów analogicznych do V można znaleźć niemało. Zauważmy mimochodem, że można zamiast hipotezy $H_5 = H_1$ przyjąć następującą, którą nazwijmy H_5': jakości a_j kolejnych partyj są takie, że plan P_0 obliczony z (14) minimizuje średnią stratę. Ta matematycznie skomplikowana hipoteza nie utrudnia obliczenia planu P_0, a wcale nie jest dalsza od rzeczywistości niż inne.

§ 7. Porównanie metod $\mathcal{R}$ i $\mathcal{P}$. Metody te nie podpadają bezpośrednio pod schemat §§ 5 i 6, a to dlatego, że nie operują pojęciem całkowitej straty: w praktyce kalkulacja szkody z decyzji jest zwykle trudna i dlatego (niezupełnie świadomie) przyjął się postulat 5-procentowego ryzyka błędnej decyzji, który się narzuca wszystkim planom. Wyraźnie mówiąc, ustala się z góry (zależnie od rodzaju towaru) dwie jakości, a_1 i a_2 $(0 < a_1 < a_2 < 1)$, i uważa za złe partie o jakościach $\leqslant a_1$, a za dobre partie o jakościach $\geqslant a_2$. Decyzja jest błędna, jeżeli odrzucono dobrą partię (błąd pierwszego rodzaju) lub przyjęto złą (błąd drugiego rodzaju). Żąda się, by maksymalne prawdopodobieństwo błędu pierwszego rodzaju było 5% i maksymalne prawdopodobieństwo błędu drugiego rodzaju było 5%. Są to postulaty ograniczające swobodę wyboru planów. Postulaty te są jednak dwuznaczne, gdyż możemy prawdopodobieństwo błędu pierwszego rodzaju interpretować jako prawdopodobieństwo, że partia odrzucona będzie dobra lub że partia dobra zostanie odrzucona; ta dwuznaczność odnosi się oczywiście także do błędu drugiego rodzaju. Pierwsza interpretacja daje metodę $\mathcal{R}$, druga zaś metodę $\mathcal{P}$ (p. wstęp).

Dla ogólności i dla umożliwienia porównania obu metod założymy, że plany odbiorcze podpadają pod ogólny schemat opisany w § 3. Przyjmiemy także zasadę minimalnej straty gospodarczej (por. § 1), która jednak tutaj będzie identyczna z kosztem badania. Jest to zgodne z programem analizy sekwencyjnej Walda, którego dezyderatem jest zmniejszenie ilości sztuk próbnych do minimum; oczywiście dezyderat ten przyjmiemy dla obu metod. Jak zauważyliśmy, szkodę z decyzji się pomija, więc funkcję S z § 4 zastąpimy przez $s(P, a, t)$ wyrażającą koszt badania.

Sformalizujmy metodę $\mathcal{R}$. Usuwamy a przez całkowanie. Ma być

$$\int_0^1\int_0^1 s(P, a, t)\,da\,dt = \text{minimum} \quad \text{dla} \quad P = P_0, \tag{15}$$

przy czym dopuszcza się tylko plany P spełniające postulat 5% ryzyka, to znaczy tylko takie, które z prawdopodobieństwem 1 w ciągu kolejnych partyj wykażą wśród odrzuconych co najwyżej 5% dobrych, a wśród przyjętych co najwyżej 5% złych. Plan wyznaczony przez (15) zrealizuje minimum średniego kosztu badania, gdy przyjmiemy hipotezę H_1 (§ 5, I). Dowód nie różni się od dedukcji podanej w § 5, punkt I.

Sformalizujmy metodę $\mathcal{P}$. Usuwamy a przez ustalenie i dodawanie. Ma być

$$(16) \qquad \int_0^1 s(P, a_1, t)\,dt + \int_0^1 s(P, a_2, t)\,dt = \text{minimum} \qquad \text{dla} \quad P = P_0,$$

przy czym dopuszcza się tylko plany P, które z prawdopodobieństwem 1 wykażą w ciągu kolejnych partyj wśród dobrych co najwyżej 5% odrzuconych, a wśród złych co najwyżej 5% przyjętych. Plan wyznaczony przez (16) zrealizuje minimum średniego kosztu badania, gdy przyjmiemy hipotezę H_4 (§ 5, IV). Dowód nie różni się od wywodów § 5, punkt IV.

Zauważmy, że jest różnica między przykładami z § 5 a tu podanymi dwiema metodami. Tam można było mówić o planach, które zachowywały sens, nawet gdy odpowiednie hipotezy nie były spełnione. Tu spełnienie dezyderatu (15) *wraz z postulatem* 5% ryzyka jest zadaniem nierozwiązalnym bez *jakiejś* hipotezy co do rozkładu jakości w kolejnych partiach. To samo odnosi się do (16). Te dowolne hipotezy nadają rozwiązalność zadaniom. Spośród nich H_1 dla (15), H_4 dla (16) realizują optyma gospodarcze.

Wykazaliśmy zupełny paralelizm metod $\mathcal{R}$ i $\mathcal{P}$. Pospolite mniemanie, jakoby pierwsza z nich operowała jakąś arbitralną hipotezą, druga zaś była wolna od hipotez, polega na złudzeniu. Pokazaliśmy, że istnieją dwa analogiczne sformułowania problematu SKJ, które przy odpowiednich hipotezach (H_1 i H_4) prowadzą jedno do $\mathcal{R}$, drugie do $\mathcal{P}$. Gdy obliczymy odpowiednie plany (P_1 i P_4), to będą one optymalne, gdy odpowiednie hipotezy będą spełnione; jeżeli nie, to i tak plany P_1 i P_4 zachowają sens i będzie ich można używać. Dlatego upada zarzut, jakoby plan P_1 zawisał w powietrzu bez hipotezy H_1, a plan P_2 był ugruntowany niezależnie od hipotez; obydwa plany nadają się do użytku bez hipotez i obydwa tracą przywilej optymalności, gdy hipotezy tracą ważność.

Gdyby chodziło tylko o sprawę $\mathcal{R}$ *versus* $\mathcal{P}$ można by było skrócić nasze rozważania. Ale musieliśmy postawić rzecz na szerszym fundamencie z powodu jej zasadniczej wagi. Problemat SKJ jako zagadnienie postępowania, w którym interweniuje los, i to w dwojaki sposób, prowadzi do konieczności wyrugowania dwóch parametrów. Gdy zdamy sobie z tego sprawę, jasno okaże się niezbędność hipotez przy wszelkich metodach SKJ, pospolitych i niezwykłych, łatwych i skomplikowanych. Gdy obliczymy optymalne plany, to pozostaną one poprawne i bez hipotez, a zatem ich określenie staje się konwencjonalne: „będziemy używali planu P_j, który *byłby* optymalny przy hipotezie H_j".

Także w naukach biologicznych i technicznych powtarza się kontrowersja, którą tu przestudiowaliśmy na przykładzie metod $\mathcal{R}$ i $\mathcal{P}$, a tylko terminologia zmienia się zależnie od dziedziny. Zarówno w zagadnieniach lekarskich jak w kwestiach bezpieczeństwa konstrukcyj budowlanych pojawia się złudzenie, jakoby ograniczenie się do zdań matematycznie poprawnych, choćby luźno związanych z problematem właściwego postępowania, mogło nas uwolnić od sprawy hipotez i stanowiło postęp w porównaniu z metodą opartą na regule Bayesa. Postęp jest tylko w tym, że nowe metody mogą być użyteczniejsze niż stare, ale nie dlatego, iżby utajone w nich hipotezy były naturalniejsze, lecz dlatego, że czasem mogą się okazać prostszymi w użyciu; ich konwencjonalny charakter nie da się przez to zatrzeć.

Wszystko, co tu napisaliśmy, uważamy za dowód, że tak zwany paradoks Bayesa w rachunku prawdopodobieństwa nie istnieje. Wystarczyłoby wymyślić termin, który oznacza prawdopodobieństwo obliczone przy hipotezie równomiernego rozkładu *a priori*, tak jak wiarogodność oznacza prawdopodobieństwo obliczone przy pewnej innej hipotezie (którą się przemyca ukrywając ją zręcznie w tym właśnie terminie), żeby wykazać nicość rzekomego paradoksu. Ale bez konfrontacji teorii z zastosowaniami niniejsza rewizja poglądów nie dałaby się przeprowadzić, choć jej wyniki nie są związane istotnie z tymi lub innymi zastosowaniami. Autor zamierza powrócić do tej kwestii w najbliższej przyszłości.

Państwowy Instytut Matematyczny

(Praca wpłynęła dnia 8. 7. 1952 r.)

Г. ШТЕЙНГАУЗ (Вроцлав)

ОСНОВЫ СТАТИСТИЧЕСКОГО КОНТРОЛЯ КАЧЕСТВА

РЕЗЮМЕ

Статистический контроль качества приводит, на основе плана выборочного приема, к решению о партии товара в зависимости от качества екземпляров в выборке. План, который дает минимум экономического убытка, является лучшим. Убыток равен сумме: потери из-за неправильного решения и стоимости выбора; он дается в виде функции $S(\alpha, t)$, где α — качество партии, t — действительное число, связанное взаимно однозначным соответствием с множеством случайных величин характеризующих экземпляры в выборке. Для того, чтобы определить лучший план, необходимо дать числовое значение функции $S(\alpha, t)$, считая α и t явными переменными. Интегрирование по t превращает убыток S в среднее значение убытка, причем α можно исключить разными методами. Каждый из этих методов приводит к некоторому оптимальному плану, если принять соответствующую гипотезу о распределении качества $\{\alpha_n\}$ в последовательности партий, предоставленных к приему, и рассматривать — вместо убытка в одном приеме — средний убыток в множестве приемов. Эта схема охватывает все известные методы статистического контроля качества. Их экономическая рациональность зависит от рациональности гипотезы о распределении α_n. Это основывается на том, что все гипотезы, относящиеся к обычно применяемым методам, — искусственные, а иллюзия, что некоторые современные методы превосходят другие, состоит в замаскировании гипотез. Анализ пяти типичных методов показывает, что так называемый ретроспективный метод, осуждаемый из-за неправильного применения правила Bayesa, никоим образом не худший, чем другие методы. Наоборот, гипотеза, относящаяся к нему, более естественная, чем гипотезы относящиеся к другим методам.

H. STEINHAUS (Wrocław)

THE PRINCIPLES OF STATISTICAL QUALITY CONTROL

SUMMARY

The statistical quality control makes the decision concerning a lot of goods dependent upon the quality of specimens in a sample

by means of a sampling acceptance plan. A plan which minimizes the economic loss is the best. The loss is the sum of the damage resulting from a wrong decision and of the cost of sampling; it has the form $S(a,t)$, where a is the quality of the lot and t — a real number associated in a bi-unique manner with a set of random numbers indicating the specimens assigned for a sample. In order to determine the best plan it is necessary to give a numerical value to the function $S(a,t)$ by making a and t into apparent variables. Integration with respect to t changes the loss S into the expected loss, while a can be eliminated by various methods. Each of those methods leads to a certain optimum plan if we accept a suitable hypothesis concerning the distribution of the quality $\{a_n\}$ in a sequence of lots presented for acceptance, and consider — instead of the loss in a single acceptance — the mean loss in a set of acceptances. All known methods of statistical quality control fall under this scheme. Their economic rationality depends on the rationality of the hypothesis concerning the distribution of a_n. It has been found that all hypotheses appertaining to methods commonly applied are artificial, and the illusion that some modern methods are superior to others rests on the hypotheses being camouflaged. An analysis of five typical methods shows that the so called retrospective method, condemned on account of its allegedly wrong application of the Bayes rule, is by no means inferior to other methods. On the contrary, the hypothesis appertaining to it is more natural than the hypotheses appertaining to other methods.