Статистичні методи теорії ризику

© Ямненко Р.Є

КНУ імені Тараса Шевченка
механіко-математичний факультет
магістратура "Актуарна та фінансова математика"

I семестр 2012

Зміст

Математичне формулювання задачі оцінювання ризику

Індивідуальний ризик

Індивідуальний ризик можна розглядати як "чорний ящик", який продукує величини сумарних збитків X_j, $j = 1, 2, \ldots, n$, де X_j позначає величину позовів за рік j (або інший страховий період j). Прикладами "чорного ящика", пов'язаного з індивідуальним ризиком, є:

- індивідуальний водій у страхуванні цивільно-правової відповідальності;
- група застрахованих осіб у страхуванні життя;
- всі працівники фірми у страхуванні від нещасних випадків на виробництві;
- компанія-цедент у перестрахуванні.

На основі спостережень за попередні періоди $\mathbf{X} = (X_1, \ldots, X_n)$ ми хочемо визначити ризикову премію для сумарних виплат у майбутньому періоді, наприклад, X_{n+1}. Для цього треба зробити певні припущення стосовно функції розподілу в.в. X_j.

Найпростіші стандартні припущення

A1) Стаціонарність: всі в.в. X_j однаково розподілені з (умовною) функцією розподілу $F(x)$.

A2) (Умовна) незалежність: в.в. X_j, $j = 1, 2, \ldots$, є (умовно) незалежними (відносно даного розподілу $F(x)$).

Припущення стаціонарності дозволяє встановити зв'язок між минулим та майбутнім. Це припущення чи, можливо, слабше завжди потрібне для обчислення страхових премій на основі історичних даних. Надалі припущення A1) (сильної) стаціонарності буде послаблене і узагальнене.

На практиці стаціонарності можна досягти, роблячи коригування відносно інфляції, індексації, усунення тренду тощо. (Умовну) незалежність пояснимо дещо згодом.

Зазвичай у страховій практиці мають справу з такими ситуаціями:

i) F невідома;

ii) F змінюється від ризику до ризику.

Для того, щоб чіткіше формалізувати i) та ii), використаємо звичайну познаку математичної статистики: проіндексуємо F за параметром ν та будемо писати F_ν, маючи на увазі, що:

- значення ν невідоме;
- ν змінюється від ризику до ризику.

Параметризація можлива завжди. Загалом, вважатимемо ν елементом деякого абстрактного простору Θ. Будемо казати, що ν задає *профіль ризику*.

Точна індивідуальна премія

Під принципом $\mathcal{H}$ обчислення премії будемо розуміти функцію, яка ставить у відповідність дійсне число в.в. X з функцією розподілу $F(x)$, тобто

$$X \mapsto \mathcal{H}(X)$$

або, вказуючи на те, що значення $\mathcal{H}$ залежить саме від F,

$$F \mapsto \mathcal{H}(F).$$

Найбільш відомими є наступні класичні принципи обчислення стахової премії.

- *Принцип очікуваного значення:* $X \mapsto (1+\alpha)\mathsf{E}X$, $\alpha > 0$.
- *Принцип стандартного відхилу:* $X \mapsto \mathsf{E}X + \beta\sigma_X$, $\beta > 0$.
- *Принцип дисперсії:* $X \mapsto \mathsf{E}X + \gamma\mathsf{D}X$, $\gamma > 0$.
- *Експоненційний принцип:* $X \mapsto \frac{1}{\delta}\ln \mathsf{E}e^{\delta X}$, $\delta > 0$.

Для кожного з цих принципів премію $\mathcal{H}(X)$ можна розкласти на *нетто-премію* EX та додатне *навантаження на ризик*, яке визначають за допомогою параметрів α, β, γ чи δ.

Економічна потреба у навантаженні на ризик пояснюється тим фактом, що страховим компаніям потрібен ризиковий капітал для того, щоб управлятися із волатильністю страхових результатів та щоб бути спроможними виконати зобов'язання перед застрахованими у несприятливий рік.

Інвестори теж потребують компенсації за піддавання ризику свого капіталу. Навантаження на ризик компенсує цей ризик і може розглядатися як вартість ризикового капіталу.

Застосовуючи принцип обчислення ризику $\mathcal{H}$, отримана точна індивідуальна премія (для ризику з профілем ν) дорівнюватиме $\mathcal{H}(F_\nu)$. Тут ми розглядатимемо лише нетто-премії. Отже, перейдемо до визначення.

Точна індивідуальна премія

Точна індивідуальна премія(Correct individual premium) за ризиком із профілем ν дорівнює

$$P^{\text{ind}}(\nu) = \mathrm{E}(X_{n+1}|\nu) \equiv \mu(\nu). \tag{1}$$

Точну індивідуальну премію іноді називають *чесною премією*. З метою спрощення писатимемо E_ν замість $\mathrm{E}(\cdot|F_\nu)$ чи $\mathrm{E}(\cdot|\nu)$.

Задачу оцінювання індивідуального ризику можна описати як визначення величини $\mu(\nu)$. Однак у страховій практиці значення як ν, так і $\mu(\nu)$ невідомі. Отже, треба знайти оцінку $\widehat{\mu(\nu)}$.

Задача оцінювання ризику в колективі

Страхова компанія страхує багато видів ризику. Щоб оцінити, їх групують за класами “подібності”, використовуючи так звані “об’єктивні” характеристики.

У страхуванні автотранспорту прикладами таких характеристик ризику є об’єм двигуна, рік виготовлення автомобіля та відношення потужність/маса, а також такі індивідуальні характеристики, як вік водія, стать та регіон.

У страхуванні підприємств від пожеж важливими характеристиками можуть бути тип конструкції застрахованої будівлі, вид бізнесу, який у ній ведеться, чи наявні засоби пожежогасіння.

На практиці питання, як точно формуються такі групи, є важливим, але воно не є предметом нашого дослідження. Важливим для нас є те, що ми не розглядаємо кожен ризик окремо, а досліджуємо його разом із групою "подібних" ризиків, яку називають колективом.

Розглянемо наступну ситуацію: нехай кожен ризик i у колективі характеризується своїм індивідуальним профілем ν_i. Параметри ν_i – це елементи множини Θ, де Θ – множина усіх можливих значень (невідомих) профілів ризику в колективі.

Якщо колектив однорідний, то Θ складається лише з однієї точки. Це відповідає класичній точці зору страхування: кожен член колективу має такий самий ризиковий профіль, а тому і однакову функцію розподілу відповідної величини сумарних позовів.

Однак групи чи колективи, які розглядають у страхування, в більшості неоднорідні. Інакше кажучи, значення параметра ν ризиків у колективі не однакові, швидше вони є вибірками, взятими з множини Θ, що містить більше одного елемента.

Але, хоч ризики у колективі і різні, всі вони належать до одного колективу, тобто взяті з тієї самої множини Θ. Це і означає Θ.

Це і означає "подібність" ризиків у колективі.

Конкретні значення ν, які відповідають різним ризикам у колективі, зазвичай страховику невідомі. Але, маючи певну попередню ("апріорну"), зокрема статистичну, інформацію, страховик дещо знає про структуру колективу.

Наприклад, він знає, що більшість водіїв є "кращими" ризиками, бо зрідка подають позови, і лише від малої частки водіїв позови виникають часто.

Формально (принаймні теоретично) цю інформацію можна виразити через імовірнісний розподіл $U(\nu)$, заданий на параметричному просторі Θ.

Ймовірнісний розподіл $U(\nu)$ називають *структурною функцією колективу*.

Інтерпретувати $U(\nu)$ можна кількома способами.

- У частотній інтерпретації розглядають значення ν у колективі як випадкову вибірку з деякої фіксованої множини Θ. Тоді функція $U(\nu)$ описує ідеалізовані частоти ν на множині Θ. Таку інтерпретацію називають *емпірично баєсівською*.
- У *чисто баєсівській* інтерпретації функцію розподілу $U(\nu)$ розглядають як опис персональних уявлень, попереднє знання та досвід актуарія.

Визначимо тепер *колективну премію* за такою формулою:

$$P^{\mathrm{coll}} = \int_{\Theta} \mu(\nu) \mathrm{d}U(\nu) \equiv \mu_0. \quad (2)$$

Ще раз придивимося до індивідуальної та колективної премій.

"Точна" індивідуальна премія дорівнює

$$P^{\mathrm{ind}}(\nu) = \mu(\nu) = \mathsf{E}_{\nu} X_{n+1},$$

тобто очікуваній величині позову за індивідуальним ризиком із профілем ν за оцінюваний період $n+1$.

Оскільки ν невідоме страховику, значення $\mu(\nu)$ теж невідоме. Щоб оцінити, страховик має інформацію про надходження позовів за цим ризиком за кілька попередніх періодів. Але часто вона досить обмежена і має мале прогнозне значення.

Колективна премія

$$P^{\text{coll}} = \mu_0 = \int_{\Theta} \mu(\nu) \mathrm{d} U(\nu)$$

дорівнює середній очікуваній величині індивідуального позову за всіма ризиками в колективі.

У більшості випадків це значення теж невідоме страховику. Однак, на відміну від індивідуальної премії, у колективах нормального розміру цю премію можна оцінити зі значною точністю на основі минулих спостережень.

Для страхової компанії здатність обчислювати колективну премію (– рівень тарифу) за важливістю займає центральне місце. Якщо вимагати однакову премію P^{coll} від кожного члена колективу, рахунки будуть збалансованими: зібрані премії = сумарній величині позовів (очікуваному значенню).

Виникає питання, навіщо страхова компанія зацікавлена у точній індивідуальній премії. Відповідь проста:

найбільш конкурентним тарифом є точна індивідуальна премія.

Саме конкуренція змушує компанії пропонувати якомога чесніші тарифи. Коли компанія A встановлює тарифи на одному рівні для всіх ризиків у неоднорідному колективі, кращі ризики платитимуть забагато, а гірші – замало. Якщо конкуруюча компанія B запропонує тарифи, які більш диференційовані та чесні, то її премії для кращих ризиків будуть дешевшими. У порівнянні з A, компанія B привабливіша для кращих ризиків, але не для гірших. Це може призвести до катастрофічних наслідків для A: вона втратить кращі ризики і збільшить кількість гірших.

Таку ситуацію називають антивідбором. Колектив змінився і структурна функція страховика стала менш сприятливою.

Баєсівська статистика

Математично проблему оцінювання колективного ризику можна описати у термінах баєсівської статистики.

Кожен ризик характеризується своїм індивідуальним профілем ν, який у свою чергу є реалізацією в.в. W, заданої на параметричному просторі Θ, причому

i) за умови події $W = \nu$ в.в. $X_1, X_2, \dots$ н.о.р. з функцією розподілу F_ν;

ii) W є в.в. з функцією розподілу U.

У цій моделі індивідуальна премія теж стає в.в. $\mu(W)$. Ми не знаємо її точного значення, але дещо знаємо про можливі значення $\mu(W)$ та з якою ймовірністю вони здійснюються. Тому природно розглядати $\mu(W)$ як в.в.

Визначимо *індивідуальну премію* через умовне математичне сподівання
$$P^{\text{ind}} = \mu(W) = \mathsf{E}(X_{n+1}|W),$$
тому дійсно P^{ind} – в.в.

Колективна премія
$$P^{\text{coll}} = \mu_0 = \int_{\Theta} \mu(\nu)\mathrm{d}U(\nu) = \mathsf{E}X_{n+1},$$
є безумовним мат. сподівання – детермінованою величиною.

За відсутності будь-якої попередньої інформації всі ризики рівні.

Якщо ми знаємо лише те, що у портфелі є кращі та гірші ризики, то не можемо апріорі бачити, до якого класу належить ризик. Лише апостеріорі, після зроблених спостережень за індивідуальним ризиком можна робити висновки.

Це зауваження формалізує наші інтуїтивні уявлення про колектив, як групу різних, проте подібних ризиків.

Повертаючись до припущення A2), зауважимо, що в.в. $X_1, X_2, \ldots$ лише умовно незалежні при заданому значенні W. Коли розглядати безумовно, вони додатно корельовані. Дійсно,

$$\mathrm{cov}(X_1, X_2) = \mathsf{E}(\mathrm{cov}(X_1, X_2)|W) + \mathrm{cov}(\mathsf{E}(X_1|W), \mathsf{E}(X_2|W))$$

$$= \mathrm{cov}(\mu(W), \mu(W)) = \mathsf{D}\mu(W).$$

Для кожного ризику треба оцінити точну премію $\mu(W)$ якомога точніше.

Однією з потенційних оцінок є колективна премія μ_0, тобто премію для конкретного розглянутого ризику оцінюють за допомогою усереднення очікуваних значень за всім колективом. Таку оцінку можна використовувати для нового ризику, у якого ще немає набутої історії позовів.

Однак ця оцінка не бере до уваги індивідуальний досвід. Якщо ризик спостерігався протягом n років, а $\mathbf{X}$ позначає вектор сумарних величин позовів за цей період, то необхідно долучити цю інформацію до процесу оцінювання.

Таким чином, виникає премія *найкращого набутого досвіду*, яка ґрунтується на векторі $\mathbf{X}$ історії індивідуальних позовів.

Її називають *баєсівською премію* і визначають за формулою:

$$P^* = \widetilde{\mu(W)} = \mathrm{E}(\mu(W)|\mathbf{X}). \tag{3}$$

Приклад

Розглянемо індивідуальний ризик, який може спричинити за один рік загальну величину позовів, рівну або 10 000, або 0 (позовів не було).

Цей індивідуальний ризик належить до колективу, який містить три типи ризиків: кращі (65%), помірні (30%) та гірші (5%).

Умовний розподіл задано таблицею

Загальна сума позовів	Умовні ймовірності		
	Кращі	Помірні	Гірші
0	97%	95%	90%
10 000	3%	5%	10%

i) Обчисліть P^{ind} та P^{coll}.

ii) Припустимо, що страховик спостерігав за індивідуальним ризиком протягом одного року. Знайдіть P^* для всіх можливих спостережень (сумарної величини позовів 0 чи 10 000).

iii) Обчисліть P^* для всіх можливих спостережень у разі, коли наявний дворічний досвід за позовами та за звичайного припущення про умовну незалежність позовів між двома роками за умови даного параметру ризику W.

Розв'язок

i) Нехай X – сумарна величина позовів за рік. Позначимо можливі значення параметру ризику W: кращі, помірні, гірші – через ν_1, ν_2, ν_3 відповідно. Тоді за формулою (1)

$$P^{\text{ind}}(\nu_1) = 0 \cdot \mathrm{P}(X = 0 | W = \nu_1) + 10\,000 \cdot \mathrm{P}(X = 10\,000 | W = \nu_1) = 300;$$

$$P^{\text{ind}}(\nu_2) = 500; \quad P^{\text{ind}}(\nu_3) = 1\,000.$$

Оскільки структурна функція колективу $U(\nu) = \mathrm{P}(W = \nu)$ має вигляд: $U(\nu_1) = 0,65$, $U(\nu_2) = 0,3$, $U(\nu_3) = 0,05$, то за формулою (2) знаходимо значення колективної премії:

$$\mathrm{P}^{\text{coll}} = \sum_{i=1}^{3} P^{\text{ind}}(\nu_i) U(\nu_i) = 395.$$

ii) Використовуючи формулу Баєса (**??**), обчислимо апостеріорні ймовірності

$$U(\nu|x) = \frac{F_\nu(x)U(\nu)}{\sum\limits_{i=1}^{3} F_{\nu_i}(x)U(\nu_i)}$$

та $P^* = \mathrm{E}(P^{\mathrm{ind}}|X = x) = \sum\limits_{i=1}^{3} P^{\mathrm{ind}}(\nu_i)U(\nu_i|x)$. Маємо

	W	ν_1	ν_2	ν_3	
x	P^{ind}	300	500	1 000	P^*
0	$U(\nu\|x)$	0,6564	0,2967	0,469	392,14
10 000	$U(\nu\|x)$	0,4937	0,3797	0,1266	464,56

iii) Спостереження $\mathbf{X} = (X_1, X_2)$ тепер є двовимірним вектором із можливими значеннями $\mathbf{x_1} = (0, 0)$, $\mathbf{x_2} = (10\,000, 0)$, $\mathbf{x_3} = (0, 10\,000)$, $\mathbf{x_4} = (10\,000, 10\,000)$.

З умовної незалежності позовів випливає, що $F_\nu(\mathbf{x}) = F_\nu(x_1, x_2) = F_\nu(x_1)F_\nu(x_2)$. Тоді знову за формулою Баєса (**??**) обчислимо апостеріорні ймовірності

$$U(\nu|\mathbf{x}) = \frac{F_\nu(\mathbf{x})U(\nu)}{\sum\limits_{i=1}^{3} F_{\nu_i}(\mathbf{x})U(\nu_i)}$$

та $P^* = \mathsf{E}(P^{\mathrm{ind}}|\mathbf{X} = \mathbf{x}) = \sum\limits_{i=1}^{3} P^{\mathrm{ind}}(\nu_i)U(\nu_i|\mathbf{x})$.

Маємо

$\mathbf{x}$	W / P^{ind}	ν_1 / 300	ν_2 / 500	ν_3 / 1 000	P^*
$\mathbf{x_1}$	$U(\nu\|\mathbf{x})$	0,6627	0,2934	0,439	389,40
$\mathbf{x_2}$	$U(\nu\|\mathbf{x})$	0,5022	0,3783	0,1195	459,30
$\mathbf{x_3}$	$U(\nu\|\mathbf{x})$	0,5022	0,3783	0,1195	459,30
$\mathbf{x_4}$	$U(\nu\|\mathbf{x})$	0,3188	0,4087	0,2725	572,48

Для більш глибокого дослідження властивостей баєсівської премії скористаємося апаратом статистичної теорії прийняття рішень. Вона допоможе зрозуміти, в якому сенсі премія P^* є "найкращою" і чому її називають баєсівською.

Елементи статистичної теорії прийняття рішень

Розглянемо експеримент, можливі результати якого ν належать простору Θ. Припустимо, що статистик, не знаючи ще результату експерименту, приймає рішення, наслідки якого залежать від результату експерименту.

Нехай D позначає простір усіх можливих рішень d, які може прийняти статистик, а функція $L(\nu, d) \colon \Theta \times D \mapsto \mathbb{R}$ відображає збитки статистики від прийняття рішення d, коли значення невідомого результата експерименту W набуває значення ν.

Будемо вважати, що заданий (фіксований) розподіл ймовірностей U на просторі результатів експерименту Θ, тобто U – вимірна функція відносно деякої σ-алгебри $\mathcal{F}$ підмножин Θ.

Для будь-якого фіксованого рішення $d \in D$ функція L визначає розподіл ймовірностей U_d. Зокрема, для будь-якого $x \in \mathbb{R}$ значення $U_d(x)$ визначається наступним чином:

$$U_d(x) = \mathrm{P}(L(W, d) < x) = \mathrm{P}(\{\nu \in \Theta\colon L(\nu, d) < x\}). \quad (4)$$

Для того, щоб розподіл U_d з (4) був коректно визначений, необхідно, щоб виконувалась така умова: для довільного $x \in \mathbb{R}$ множина $\{\nu \in \Theta\colon L(\nu, d) < x\}$ належить σ-алгебрі $\mathcal{F}$.

Припустимо, що ця умова виконана для будь-якого рішення $d \in D$. Тоді для кожного розподілу ймовірностей U_d, для якого функція L інтегровна, середня збитковість

$$\mathrm{E}(L|U_d) = \int\limits_{\Theta} L(\nu, d) \mathrm{d} U(\nu). \quad (5)$$

Статистик повинен вибирати рішення d, що мінімізує $\mathrm{E}(L|U_d)$. Коли рішення приймають без інформації про результат W експерименту, W називають параметром, а множину Θ можливих значень W – параметричним простором.

Отже, задача прийняття рішень визначається параметричним простором Θ, простором рішень D і дійснозначною функцією втрат L, яка визначена на добутку $\Theta \times D$. При будь-якому $(\nu, d) \in \Theta \times D$ число $L(\nu, d)$ є збиток статистика від прийняття рішення d, коли значення параметра W дорівнює ν. Передбачається, що $L(\cdot, d)$ є $\mathcal{F}$-вимірною функцією на просторі Θ при всіх $d \in D$.

Нехай параметр W – це в.в. із заданим розподілом імовірностей $U(\nu)$, $\nu \in \Theta$. Для довільного рішення $d \in D$ середній збиток $R(d)$, який називають *ризиком*, визначають за формулою

$$R(d) = R(d, U) = \int\limits_{\Theta} L\,(\nu, d)\,\mathrm{d}U(\nu). \tag{6}$$

Ми будемо вважати, що інтеграл у (6) скінченний для всіх $d \in D$. Рішення d, для яких це припущення не виконане, як правило, можуть бути виключені з множини D.

Зі співвідношення (5) випливає, що статистик вибирає рішення d, яке мінімізує ризик $R(d)$.

Баєсівський ризик та баєсівські рішення

Розглянемо задачу прийняття рішення з параметричним простором Θ, простором рішень D і функцією втрат $L(\nu, d)$. Для будь-якого розподілу U значень параметра W *баєсівський ризик* $R^* = R^*(U)$ визначають як точну нижню границю ризиків $R(d)$, коли рішення $d \in D$, тобто

$$R^* = \inf_{d \in D} R(d) . \tag{7}$$

Рішення d^*, ризик від якого дорівнює баєсівському ризику, називають *баєсівським рішенням* за розподілу U. Отже, рішення d^* називають баєсівським за розподілу U тоді і лише тоді, коли $R(d^*) = R^*$.

Якщо U є розподілом параметра W, то будь-яке баєсівське рішення за U буде оптимальним для статистика, оскільки ні при якому іншому рішенні ризик не може бути меншим.

Можливо, однак, що жодне рішення з класу D не буде баєсівським. Ця ситуація реалізується в тому випадку, коли нижня границя у (7) не досягається на множині D. У цьому випадку статистику варто вибирати рішення $d \in D$, для якого ризик $R(d)$ досить мало відрізняється від баєсівського.

Оскільки ці труднощі не є основними ні в теорії, ні в практиці прийняття рішень, ми будемо, як правило, припускати надалі, що для всіх розподілів U баєсівський ризик R^* досягається за деякого рішення $d \in D$.

Приклад

Розглянемо задачу прийняття рішень, у якій параметричний простір $\Theta = \{\nu_1, \nu_2, \nu_3, \nu_4\}$, простір рішень $D = \{d_1, d_2, d_3\}$, а функція втрат $L(\nu, d)$ задана таблицею 1 нижче. Припустимо, що розподіл ймовірностей $U(\nu) = \mathrm{P}(W = \nu)$ параметра W такий: $U(\nu_1) = \frac{1}{8}$, $U(\nu_2) = \frac{3}{8}$, $U(\nu_3) = \frac{1}{4}$ та $U(\nu_4) = \frac{1}{4}$. Знайдемо баєсівське рішення за такого розподілу U.

Табл.1

	d_1	d_2	d_3
ν_1	0	2	3
ν_2	1	0	2
ν_3	3	4	0
ν_4	1	2	0

Розв'язок

Обчислимо ризик при заданому розподілі ймовірностей:

$$R(d) = \int_{\Theta} L(\nu, d)\, \mathrm{d}U(\nu) = \sum_{i=1}^{4} L(\nu_i, d)\, U(\nu_i)$$

$$= \frac{1}{8} L(\nu_1, d) + \frac{3}{8} L(\nu_2, d) + \frac{1}{4} L(\nu_3, d) + \frac{1}{4} L(\nu_4, d).$$

Враховуючи значення функції втрат $L(\nu, d)$, отримаємо

$$\begin{aligned} R(d_1) &= \frac{1}{8}(0+3+6+2) = \frac{11}{8}; \\ R(d_2) &= \frac{1}{8}(2+0+8+4) = \frac{14}{8}; \\ R(d_3) &= \frac{1}{8}(3+6+0+0) = \frac{9}{8}. \end{aligned}$$

Звідси визначаємо баєсівський ризик:

$$R^* = \inf_{d \in D} R(d) = R(d_3) = \frac{9}{8}.$$

Отже, $d^* = d_3$ – баєсівське рішення.

Приклад

Нехай параметричний простір Θ складається з двох точок 0 і 1, а простір рішень D складається з усіх чисел d інтервалу $0 \le d \le 1$. Нехай функція втрат $L(\nu, d)$, $\nu \in \Theta$, $d \in D$, визначається за

$$L(\nu, d) = |\nu - d|^{\alpha}, \tag{8}$$

де $\alpha \ge 1$ – задане число. Нехай розподіл ймовірностей $U(\nu) = \mathrm{P}(W = \nu)$ параметра W такий: $U(0) = \frac{3}{4}$, $U(1) = \frac{1}{4}$. Знайти баєсівське рішення за такого розподілу U.

Розв'язок

а) Розглянемо спочатку випадок, коли параметр $\alpha = 1$. Тоді для всякого рішення $d \in D$ ризик $R(d)$ задається формулою

$$R(d) = L(0, d)\,\mathrm{P}(W = 0) + L(1, d)\,\mathrm{P}(W = 1)$$

$$= \frac{3}{4}d + \frac{1}{4}(1 - d) = \frac{d}{2} + \frac{1}{4}.$$

Як бачимо, що $R(d)$ буде мінімальним при $d = 0$. Отже, $d^* = 0$ – єдине баєсівське рішення і баєсівський ризик R^* дорівнює $1/4$.

Зауважимо, що коли простір рішень D визначений як напіввідкритий інтервал $0 < d \leq 1$, то баєсівський ризик, як і раніше, дорівнює $1/4$, проте жодне рішення з D не буде баєсівським.

b) Розглянемо тепер задачу за умови, що параметр $\alpha > 1$. Тоді для кожного рішення $d \in D$

$$R\left(d\right) = \frac{3}{4} d^{\alpha} + \frac{1}{4}\left(1 - d\right)^{\alpha}.$$

Значення d, яке мінімізує ризик, знаходиться диференціюванням функції $R(d)$. А саме, єдине рішення d^* має вигляд

$$d^* = \left(1 + 3^{\frac{1}{\alpha - 1}}\right)^{-1}.$$

Процедури прийняття рішень в умовах невизначеності

Розглянемо інші критерії, які найчастіше використовують на практиці під час прийняття рішень в умовах невизначеності, тобто коли структурна функція $U(\nu)$ невідома. Критерії:

- критерій Лапласа;
- мінімаксний критерій (Вальда);
- критерій Севіджа;
- критерій Гурвиця.

Ми розглядаємо задачі прийняття рішень в умовах невизначеності, коли вибір рішення з множини допустимих рішень здійснюється однією особою. Специфічною особливістю цих задач є відсутність у "особи, що приймає рішення" розумного супротивника. У тому випадку, коли в ролі супротивника виступає "природа", немає підстав припускати, що вона прагне завдати шкоди "особі, що приймає рішення".

Інформацію, необхідну для прийняття рішень в умовах невизначеності, звичайно подано у формі матриці, в якій j-й стовпець відповідає рішенню d_j з множини допустимих рішень $D = \{d_1, \ldots, d_m\}$, а i-й рядок відповідає стану ν_i системи, що вивчається, з множиною можливих станів $\Theta = \{\nu_1, \ldots, \nu_k\}$.

Таким чином, якщо множина $D = \{d_1, \ldots, d_m\}$ допустимих рішень складається з m елементів, а система може знаходитися в будь-якому з $\Omega = \{\nu_1, \ldots, \nu_k\}$ можливих станів, то матриця

$$L = \{L(\nu_i, d_j)\}_{i=1,\ldots,k}^{j=1,\ldots,m}$$

є матрицею початкових даних для прийняття рішень в умовах невизначеності. Цю матрицю називають *матрицею втрат* або *матрицею витрат*.

Критерій Лапласа

Для обґрунтування цього критерію скористаємося наступними міркуваннями, що відображають основну суть *принципу недостатнього обґрунтування*.

Оскільки ймовірність перебування системи, що вивчається, в кожному її можливому стані $\nu_1, \ldots, \nu_k$ невідомі, то відсутня і необхідна інформація для висновку про те, що ця ймовірність різна. Інакше ми могли б застосувати баєсівський підхід до прийняття рішень.

Тому ми можемо припустити, що ймовірності реалізації будь-яких можливих станів системи рівні.

Таким чином, задачу можна розглядати як задачу прийняття рішень за умови, що вибирають рішення, яке забезпечує найменші втрати, тобто

$$R(d^*) = \inf_{d_j \in D} \frac{1}{k} \sum_{i=1}^{k} L(\nu_i, d_j).$$

Сформульований критерій називають *критерієм Лапласа*.

Приклад

Підприємство має визначити рівень пропозиції послуг так, щоб задовольнити потреби клієнтів протягом майбутніх свят. За попередніми прогнозами число клієнтів може набути одного з наступних значень: $\nu_1 = 200$, $\nu_2 = 250$, $\nu_3 = 300$, $\nu_4 = 350$. Для кожного з цих можливих значень існує найкращий з погляду можливих витрат рівень пропозицій d_j, а сукупність цих рівнів утворює множину D

Відхилення від рівнів d_j приводять до додаткових витрат або через неповне задоволення попиту, або через перевищення пропозиції над попитом. Матриця витрат в умовних грошових одиницях подана нижче:

$$L = \begin{pmatrix} 5 & 8 & 21 & 30 \\ 10 & 7 & 18 & 22 \\ 18 & 8 & 12 & 19 \\ 25 & 23 & 21 & 15 \end{pmatrix},$$

де L_{ij} – це витрати за $W = \nu_i$ та $d = d_j$. Знайти оптимальне рішення d^* за критерієм Лапласа.

Розв'язок

Для матриці L маємо

$$R(d_j) = \frac{1}{4}\sum_{i=1}^{4} L\left(\nu_i, d_j\right).$$

Тому

$$R(d_1) = \frac{1}{4}(5 + 10 + 18 + 25) = 14,5;$$

$$R(d_2) = 11,5; \quad R(d_3) = 18; \quad R(d_4) = 24,5.$$

Звідси маємо

$$R(d^*) = \inf_{d_j \in D} \frac{1}{4}\sum_{i=1}^{4} L\left(\nu_i, d_j\right) = 11,5.$$

Найкращим рівнем пропозиції за критерієм Лапласа буде d_2.

Мінімаксний критерій

Цей критерій є найбільш "обережним", оскільки його реалізація передбачає вибір найкращої з найгірших можливостей.

Нехай $D = \{d_1, \ldots, d_m\}$ – множина допустимих рішень, а $\Theta = \{\nu_1, \ldots, \nu_k\}$ – множина можливих станів системи, що вивчається. Якщо $L(\nu_i, d_j)$ – втрати "особи, що приймає рішення" внаслідок вибору рішення $d_j \in D$ та реалізації системою можливого стану $\nu_i \in \Theta$, то незалежно від можливих станів найбільші втрати дорівнюватимуть

$$\max_{\nu_i \in \Theta} L(\nu_i, d_j), \quad j = 1, \ldots, m.$$

За мінімаксним критерієм вибирають таке рішення d^*, що

$$R(d^*) = \min_{d_j \in D} \max_{\nu_i \in \Theta} L(\nu_i, d_j).$$

Приклад

Повернемося до попереднього прикладу. Оскільки в цьому випадку L_{ij} відображають втрати, то скористаємося мінімаксним критерієм. Для кожного допустимого вирішення $d_j \in D$ знайдемо максимальні витрати:

$$\max_{\nu_i \in \Theta} L\left(\nu_i, d_1\right) = L(\nu_4, d_1) = 25; \max_{\nu_i \in \Theta} L\left(\nu_i, d_2\right) = L(\nu_4, d_2) = 23;$$

$$\max_{\nu_i \in \Theta} L\left(\nu_i, d_3\right) = L(\nu_1, d_3) = L(\nu_4, d_3) = 21;$$

$$\max_{\nu_i \in \Theta} L\left(\nu_i, d_4\right) = L(\nu_1, d_4) = 30.$$

Потім з-поміж обчислених значень виберемо мінімальне:

$$\min_{d_j \in D} \max_{\nu_i \in \Theta} L\left(\nu_i, d_j\right) = L(\nu_1, d_3) = L(\nu_4, d_3) = 21.$$

Отже, оптимальним мінімаксним рішенням є d_3.

Критерій Севіджа

Мінімаксний критерій є настільки “песимістичним”, що може призводити до нелогічних висновків. Необхідність використання менш “песимістичного” критерію звичайно ілюструють задачею прийняття рішень в умовах невизначеності з матрицею втрат

$$L = \begin{pmatrix} 11\,000 & 10\,000 \\ 90 & 10\,000 \end{pmatrix}.$$

Застосування мінімаксного критерію приводить до вибору рішення d_2 та втрат 10 000 у разі реалізації системою кожного з можливих станів ν_1 або ν_2. Проте інтуїтивно напрошується висновок про доцільність вибору рішення d_1, оскільки не виключається можливість реалізації стану ν_2 та втрат $L(\nu_2, d_1) = 90$.

Для усунення відзначеного недоліку мінімаксного критерію замість величини $L(\nu_i, d_j)$, що характеризує втрати під час прийняття рішення d_j внаслідок реалізації стану ν_i, введемо величину

$$r(\nu_i, d_j) = L(\nu_i, d_j) - \min_{d_j \in D} L(\nu_i, d_j).$$

Фактично величина $r(\nu_i, d_j)$ відображає величину "розчарування" особи, що приймає рішення, з приводу того, що вона не вибрала найкраще рішення відносно стану ν_i системи. Тому матрицю $R = \{r(\nu_i, d_j)\}_{i=1,\dots,k}^{j=1,\dots,m}$ називають *матрицею розчарування*, а мінімаксний критерій відносно цієї матриці називають *критерієм Севіджа*. За цього критерію рішення вибирають з умови

$$R(d^*) = \min_{d_j \in D} \max_{\nu_i \in \Theta} r(\nu_i, d_j).$$

Зокрема, у попередньому прикладі, розв'язок якого з використанням мінімаксного критерію призводив до нелогічного висновку, маємо

$$\min_{d_j \in D} L(\nu_1, d_j) = 10\,000, \quad \min_{d_j \in D} L(\nu_2, d_j) = 90,$$

матриця розчарування має вигляд

$$R = \begin{pmatrix} 1\,000 & 0 \\ 0 & 9\,910 \end{pmatrix}.$$

Таким чином,

$$\max_{\nu_i \in \Theta} r(\nu_i, d_1) = r(\nu_1, d_1) = 1\,000; \quad \max_{\nu_i \in \Theta} r(\nu_i, d_2) = r(\nu_2, d_2) = 9\,910;$$

і за критерієм Севіджа оптимальним є рішення d_1. Відзначимо, що такий самий результат ми отримаємо і за критерієм Лапласа.

Приклад

Повернемося до прикладу і запишемо матрицю розчарування

$$R = \begin{pmatrix} 0 & 3 & 16 & 10 \\ 3 & 0 & 11 & 15 \\ 10 & 0 & 4 & 11 \\ 10 & 8 & 6 & 0 \end{pmatrix}.$$

Звідси отримуємо

$$\max_{\nu_i \in \Theta} r(\nu_i, d_1) = r(\nu_3, d_1) = 10; \quad \max_{\nu_i \in \Theta} r(\omega_i, d_2) = r(\nu_4, d_2) = 8;$$

$$\max_{\nu_i \in \Theta} r(\nu_i, d_3) = r(\nu_1, d_3) = 16; \quad \max_{\nu_i \in \Theta} r(\omega_i, d_4) = r(\nu_1, d_4) = 25;$$

$$\min_{d_j \in D} \max_{\nu_i \in \Theta} r(\nu_i, d_j) = r(\nu_4, d_2) = 8,$$

і оптимальним за критерієм Севіджа є рішення d_2.

Критерій Гурвиця

Цей критерій охоплює ряд підходів до прийняття рішень в умовах невизначеності від найбільш песимістичного до найбільш оптимістичного. Якщо $\{L(\nu_i, d_j)\}_{i=1,\dots,k, j=1,\dots,m}$ – матриця втрат, то найбільш оптимістичному підходу відповідає критерій, що реалізує

$$\min_{d_j \in D} \min_{\nu_i \in \Theta} L(\nu_i, d_j),$$

а найбільш песимістичному підходу відповідає критерій, що реалізує

$$\min_{d_j \in D} \max_{\nu_i \in \Theta} L(\nu_i, d_j).$$

Критерій Гурвиця встановлює баланс між найбільш оптимістичним і найбільш песимістичним підходами за допомогою середнього зваженого значення обох варіантів прийняття рішень в умовах невизначеності з вагами α та $1-\alpha$, де $0 \leq \alpha \leq 1$. Це означає, що за критерію Гурвиця вибирають рішення $d^* \in D$, що забезпечує

$$\min_{d_j \in D} \left(\alpha \min_{\nu_i \in \Theta} L(\nu_i, d_j) + (1-\alpha) \max_{\nu_i \in \Theta} L(\nu_i, d_j) \right).$$

Параметр $\alpha \in [0,1]$ називають *показником оптимізму*. Його значення вибирається особою, що приймає рішення, залежно від досвіду прийняття рішень в умовах невизначеності і особистих схильностей до оптимізму ($\alpha \to 1$) чи песимізму ($\alpha \to 0$). За відсутності яскраво виражених схильностей значення $\alpha = 0,5$ видається найбільш розумним.

Приклад

d_j	$\min\limits_i L(i,j)$	$\max\limits_i L(i,j)$	$\alpha \min\limits_i L(i,j) + (1-\alpha) \max\limits_i L(i,j)$
d_1	5	25	15
d_2	7	23	15
d_3	12	21	16,5
d_4	15	30	22,5

Скористаємося критерієм Гурвиця для розв'язку задачі з нашого прикладу, вважаючи $\alpha = 0,5$. Результати розрахунків подані в таблиці. Згідно з результатами розрахунків, оптимальне значення за критерієм Гурвиця рівне 15 і забезпечується допустимими рішеннями d_1 і d_2.

Задачі зі спостереженнями

Розглянемо задачу прийняття рішень, у якій статистик перед тим, як вибрати рішення з множини D, спостерігає значення випадкової величини X чи випадкового вектора $\mathbf{X} = (X_1, \dots, X_n)$, пов'язаного з параметром W. Спостереження $\mathbf{X}$ дають статистикові деяку інформацію про значення ν параметра W, яка допомагає йому прийняти більш раціональне рішення.

Припустимо, що для всіх $\nu \in \Theta$ та $\mathbf{x} \in \mathbb{R}^n$ заданий умовний розподіл вектора $\mathbf{X}$ за умови $W = \nu$:

$$F_\nu(\mathbf{x}) \equiv F_{\mathbf{X}|W=\nu}(\mathbf{x}|\nu) = \mathrm{P}(\mathbf{X} < \mathbf{x}|W = \nu).$$

Задачу такого типу називають *статистичною задачею прийняття рішень*. Основні її елементи – це параметричний простір Θ, простір рішень D, функція втрат $L(\nu, d)$ та сімейство умовних розподілів імовірностей $\{F_\nu, \nu \in \Theta\}$, що спостерігається статистиком до прийняття рішення.

Оскільки рішення статистика залежить від спостережуваного значення $\mathbf{X} = \mathbf{x}$, він, насправді, повинен вибрати вирішувальну функцію $\delta(\mathbf{x})$, яка задає для будь-якого допустимого значення $\mathbf{x} \in \mathbb{R}^n$ рішення $\delta\left(\mathbf{x}\right) \in D$.

У багатьох задачах статистику не потрібно мати усю вирішувальну функцію δ. Після того, як він довідався про результат спостереження $\mathbf{x}$, йому досить розглянути задачу вибору рішення $\delta(\mathbf{x})$.

Позначимо через Δ клас усіх вирішувальних функцій δ. За потребою будемо накладати на вирішувальні функції деякі обмеження на зразок вимірності. Для кожного значення ν параметра W ризик $R_\nu(\delta)$ від прийняття рішення, заданого вирішувальною функцією $\delta \in \Delta$, визначають за формулою

$$R_\nu(\delta) = \mathsf{E}(L(W, \delta(\mathbf{X})) \,|\, W = \nu) = \int\limits_{\mathbb{R}^n} L(\nu, \delta(\mathbf{x})) \mathrm{d}F_\nu(\mathbf{x}). \qquad (9)$$

Вважатимемо, що для всіх $\nu \in \Theta$ функція $L(\nu, \delta(\cdot))$ вимірна та інтегровна. Якщо для випадкового вектора $\mathbf{X}$ існує сумісна щільність $f_\nu(\mathbf{x}) \equiv f_{\mathbf{X}|W=\nu}(\mathbf{x}|\nu)$, то вираз (9) можна переписати так:

$$R_\nu(\delta) = \int\limits_{\mathbb{R}^n} L(\nu, \delta(\mathbf{x})) f_\nu(\mathbf{x}) \mathrm{d}\mathbf{x}. \tag{10}$$

Функцію $R_\nu(\delta)$ називають *функцією ризику*.

Баєсівським ризиком вирішувальної функції $\delta \in \Delta$ відносно попереднього розподілу $U(\nu)$ параметра W називають

$$R(\delta) = \int_\Theta R_\nu(\delta) \mathrm{d}U(\nu) = \int_\Theta \int\limits_{\mathbb{R}^n} L(\nu, \delta(\mathbf{x})) \mathrm{d}F_\nu(\mathbf{x}) \mathrm{d}U(\nu). \tag{11}$$

Нехай $\delta^* \in \Delta$ – така вирішувальна функція, що

$$R(\delta^*) = \inf_{\delta \in \Delta} R(\delta) = R^*. \tag{12}$$

Тоді δ^* називають *баєсівською вирішувальною функцією* за умови апріорного розподілу ймовірностей U, а R^*, як і раніше, називають *баєсівським ризиком*. Для кожного розподілу U параметра W статистику варто вибирати вирішувальну функцію δ, яка є баєсівською вирішувальною функцією за умови розподілу U.

Приклад

У тапас-меню місцевого іспанського ресторану відвідувачі можуть замовити блюдо з 20 смажених чилі за $5. На блюді завжди подають суміш з гострих та лагідних чилі, відрізнити які можна лише на смак.

Ресторан пропонує два типи страви: одна містить 4 гострих та 16 лагідних чилі, а інша – 8 гострих та 12 лагідних. Коли блюдо подано на стіл, офіціант дозволяє відвідувачу скуштувати одне чилі, а потім пропонує 50% знижку за правильний здогад: 4 чи 8 гострих чилі містить блюдо.

Голодний актуарій, який регулярно відвідує цей ресторан, хоче побудувати оптимальну стратегію вгадування числа гострих чилі.

(i) Перелічіть чотири можливі вирішувальні функції для актуарія.

(ii) Обчисліть значення функції ризику для двох різних блюд із чилі та кожної вирішувальної функції.

(iii) Визначте оптимальну стратегію для актуарія, використовуючи баєсівський критерій, та знайдіть середню ціну, яку він платитиме за блюдо з чилі, якщо ресторан порівну виготовляє обидва типи цієї страви.

Розв'язок

Позначимо через d_4 та d_8 рішення, які може прийняти актуарій, а саме, d_i відповідає вибору блюда з i гострими чилі.
Позначимо також значення в.в. X, які відображають скуштовані гострий чи лагідний чилі, через 0 та 1 відповідно.
Параметр W може набувати одне із двох значень ν_4 чи ν_8, залежно від того, блюдо зі скількома (i) гострими чилі приніс офіціант. Отже, $D = \{d_4, d_8\}$, $\Theta = \{\nu_4, \nu_8\}$.
(i) В залежності від спостереженого значення $x \in \{0, 1\}$ параметра W актуарій може вибрати одну з цих вирішувальних функцій:

$$
\begin{aligned}
\delta_1(0) &= d_4 \quad \text{та} \quad \delta_1(1) = d_4; \\
\delta_2(0) &= d_4 \quad \text{та} \quad \delta_2(1) = d_8; \\
\delta_3(0) &= d_8 \quad \text{та} \quad \delta_3(1) = d_4; \\
\delta_4(0) &= d_8 \quad \text{та} \quad \delta_4(1) = d_8.
\end{aligned}
$$

(ii) Для блюда з чотирма гострими чилі ймовірності скуштувати гострий та лагідний чилі дорівнюють відповідно

$$\mathsf{P}(X = 0|W = \nu_4) = 0,2, \quad \mathsf{P}(X = 1|W = \nu_4) = 0,8.$$

Для блюда з вісьмома гострими чилі маємо

$$\mathsf{P}(X = 0|W = \nu_8) = 0,4, \quad \mathsf{P}(X = 1|W = \nu_8) = 0,6.$$

Втрати актуарія $L(\nu_i, d_j) = 5\,\mathbb{I}_{\{i \neq j\}} + 2,5\,\mathbb{I}_{\{i = j\}}$ – ціна, яку він має заплатити за тарілку з чилі, тобто $\$5$ чи $\$2,5$.

Тоді функції ризику від кожної його вирішувальної функції за умови, що подано блюдо ν_4 чи ν_8, буде таким:

$$R_{\nu_4}(\delta_1) = \mathsf{E}(L\,(W, \delta_1(X))\,|W = \nu_4)$$

$$= \mathsf{P}(X = 0|W = \nu_4)L(\nu_4, \delta_1(0)) + \mathsf{P}(X = 1|W = \nu_4)L(\nu_4, \delta_1(1))$$

$$= 0,2 \cdot 2,5 + 0,8 \cdot 2,5 = 2,5;$$

$$R_{\nu_8}(\delta_1) = \mathsf{P}(X = 0|W = \nu_8)L(\nu_8, \delta_1(0)) +$$

$$+\mathsf{P}(X = 1|W = \nu_8)L(\nu_8, \delta_1(1)) = 0,4 \cdot 5 + 0,6 \cdot 5 = 5;$$

$$R_{\nu_4}(\delta_2) = \mathsf{P}(X = 0|W = \nu_4)L(\nu_4, \delta_2(0)) +$$

$$+\mathsf{P}(X = 1|W = \nu_4)L(\nu_4, \delta_2(1)) = 0,2 \cdot 2,5 + 0,8 \cdot 5 = 4,5;$$

$$R_{\nu_8}(\delta_2) = \mathsf{P}(X = 0|W = \nu_8)L(\nu_8, \delta_2(0)) +$$

$$+\mathsf{P}(X = 1|W = \nu_8)L(\nu_8, \delta_2(1)) = 0,4 \cdot 5 + 0,6 \cdot 2,5 = 3,5;$$

$$R_{\nu_4}(\delta_3) = \mathsf{P}(X = 0|W = \nu_4)L(\nu_4, \delta_3(0)) +$$

$$+\mathsf{P}(X = 1|W = \nu_4)L(\nu_4, \delta_3(1)) = 0,2 \cdot 5 + 0,8 \cdot 2,5 = 3;$$

$$R_{\nu_8}(\delta_3) = \mathsf{P}(X = 0|W = \nu_8)L(\nu_8, \delta_3(0)) +$$

$$+\mathsf{P}(X = 1|W = \nu_8)L(\nu_8, \delta_3(1)) = 0,4 \cdot 2,5 + 0,6 \cdot 5 = 4;$$

$$R_{\nu_4}(\delta_4) = \mathsf{P}(X = 0|W = \nu_4)L(\nu_4, \delta_4(0)) +$$
$$+\mathsf{P}(X = 1|W = \nu_4)L(\nu_4, \delta_4(1)) = 0,2 \cdot 5 + 0,8 \cdot 5 = 5;$$
$$R_{\nu_8}(\delta_4) = \mathsf{P}(X = 0|W = \nu_8)L(\nu_8, \delta_4(0)) +$$
$$+\mathsf{P}(X = 1|W = \nu_8)L(\nu_8, \delta_4(1)) = 0,4 \cdot 2,5 + 0,6 \cdot 2,5 = 2,5.$$

(iii) Матриця виплат гравця матиме такий вигляд:

	δ_1	δ_2	δ_3	δ_4
ν_4	2,5	4,5	3	5
ν_8	5	3,5	4	2,5
Баєсівський ризик $R(\delta) = \frac{1}{2}R_{\nu_4}(\delta) + \frac{1}{2}R_{\nu_8}(\delta)$	3,75	4	3,5	3,75

Отже, баєсівською вирішувальною функцією є δ_3. За такого вибору середня ціна за блюдо з чилі становитиме $\$3,5$.

Побудова баєсівських вирішувальних функцій

Нехай для розподілу ймовірностей U параметра W потрібно знайти вирішувальну функцію δ, яка мінімізує її баєсівський ризик $R(\delta)$, визначений у формулі (11).

Ми будемо припускати, що в цьому співвідношенні можна змінити порядок інтегрування. Зокрема, ця перестановка законна для всіх розподілів ймовірностей U та всіх вирішувальних функцій δ, якщо функція втрат L невід'ємна та обмежена. Після зазначеної зміни порядку інтегрування

$$R(\delta) = \int\limits_{\mathbb{R}^n} \left(\int_{\Theta} L(\nu, \delta(\mathbf{x})) \mathrm{d} U_{\mathbf{x}}(\nu) \right) \mathrm{d} F(\mathbf{x}), \qquad (13)$$

де $U_{\mathbf{x}}(\nu) = U_{W|\mathbf{X}=\mathbf{x}}(\nu|\mathbf{x})$ – умовний розподіл параметра W за умови $\mathbf{X} = \mathbf{x}$.

Вирішувальну функцію δ, яка мінімізує цей ризик, можна визначити з умови мінімізації внутрішнього інтеграла в (13) для кожного значення $\mathbf{x} \in \mathbb{R}^n$. Іншими словами, баєсівську вирішувальну функцію δ^* за умови розподілу U можна одержати так: для кожного спостережуваного значення $\mathbf{x}$ приймаємо рішення $\delta^*(\mathbf{x}) = d^* \in D$, яке мінімізує інтеграл

$$\int_\Theta L(\nu, \delta(\mathbf{x}))\mathrm{d}U_{\mathbf{x}}(\nu) = \int_\Theta L(\nu, d)\mathrm{d}U_{\mathbf{x}}(\nu). \qquad (14)$$

Зокрема, коли існує умовна щільність $u_{\mathbf{x}}(\nu) = u_{W|\mathbf{X}=\mathbf{x}}(\nu|\mathbf{x})$ розподілу параметра W за спостережуваного значення $\mathbf{X} = \mathbf{x}$, інтеграл (14) можна записати і так:

$$\int_\Theta L(\nu, d)u_{\mathbf{x}}(\nu)\mathrm{d}\nu = \mathsf{E}\left(L\left(W, d\right)|\mathbf{X} = \mathbf{x}\right). \qquad (15)$$

Отже, рішення d^*, яке мінімізує інтеграл (14), є попросту те, якому відповідає найменший середній збиток при умовному розподілі W, коли спостереженим значенням $\mathbf{X}$ виявилося $\mathbf{x}$. Іншими словами, d^* – баєсівське рішення за умовного розподілу W, коли $\mathbf{X} = \mathbf{x}$.

У статистичних задачах прийняття рішень початковий розподіл W називають *попереднім або апріорним розподілом* параметра W, бо він задає розподіл W до спостерігання $\mathbf{X}$. Умовний розподіл W за відомого значення $\mathbf{X}$ називають *апостеріорним розподілом* W, бо він задає розподіл після спостереження значення величини $\mathbf{X}$.

Корисно уявляти собі баєсівську вирішувальну функцію наступним чином. Якщо рішення приймають без попередніх спостережень, то оптимальним є баєсівське рішення за апріорного розподілу W. Якщо ж перед прийняттям рішення спостерігають значення $\mathbf{X}$, то задача прийняття рішення для статистика, власне кажучи, та ж сама, як і в першому випадку, різниця лише в тому, що апріорний розподіл W змінився на апостеріорний. Отже, тепер оптимальним є баєсівське рішення за апостеріорного розподілу W.

З цих міркувань ясно, що рішення $\delta^*(\mathbf{x_0})$, задане баєсівською вирішувальною функцією δ^* для певного значення $\mathbf{x_0}$, яке є результатом спостереження, можна знайти і без того, щоб обчислювати рішення $\delta^*(\mathbf{x})$ для всіх значень $\mathbf{x}$. Баєсівську вирішувальну функцію δ^* для розподілу ймовірностей U можна знаходити і без обчислення баєсівського ризику R^*.

Приклад

Нехай $\Theta = \{\nu_1, \nu_2\}$, $D = \{d_1, d_2\}$, а функція втрат L задано таблицею

	d_1	d_2
ν_1	0	5
ν_2	10	0

Припустимо, що статистик може спостерігати в.в. X з такими умовним розподілом $F_\nu(x) = \mathrm{P}(X = x | W = \nu)$:

$$F_{\nu_1}(1) = 3/4,\ F_{\nu_1}(0) = 1/4,\ F_{\nu_2}(1) = 1/3,\ F_{\nu_2}(0) = 2/3. \quad (16)$$

Потрібно побудувати баєсівську вирішувальну функцію за умови, що апріорний закон розподілу параметра W такий:

$$U(\nu_1) = \mathrm{P}(W = \nu_1) = u, \quad U(\nu_2) = 1 - u, \quad u \in (0, 1). \quad (17)$$

Розв'язок

Нехай $U_x(\nu)$ – апостеріорна ймовірність події $W = \nu$, якщо спостерігалось значення x величини X, тобто

$$U_x(\nu) = \mathrm{P}(W = \nu | X = x).$$

Застосувавши формулу Баєса (**??**), матимемо

$$U_x(\nu) = \frac{\mathrm{P}(X = x|\, W = \nu)\,\mathrm{P}(W = \nu)}{\sum\limits_{i=1}^{2} \mathrm{P}(X = x|\, W = \nu_i)\,\mathrm{P}(W = \nu_i)}$$

$$= \frac{F_\nu(x)U(\nu)}{\sum\limits_{i=1}^{2} F_{\nu_i}(x)U(\nu_i)}.$$

Отже, з рівностей (16) – (17) випливає, що

$$U_1(\nu_1) = \frac{\frac{3}{4}u}{\frac{3}{4}u + \frac{1}{3}(1-u)},$$

$$U_1(\nu_2) = \frac{\frac{1}{3}(1-u)}{\frac{3}{4}u + \frac{1}{3}(1-u)},$$

$$U_0\,(\nu_1) = \frac{\frac{1}{4}u}{\frac{1}{4}u + \frac{2}{3}(1-u)},$$

$$U_0\,(\nu_2) = \frac{\frac{2}{3}(1-u)}{\frac{1}{4}u + \frac{2}{3}(1-u)}.$$

Після спостереження значення x в.в. X треба вибрати одне з рішень: d_1 чи d_2. З таблиці видно, що ризик від прийняття рішення d_1 дорівнює

$$R(d_1) = 10U_x(\nu_2) = 10(1 - U_x(\nu_1)),$$

а ризик від прийняття d_2 дорівнює

$$R(d_2) = 5U_x(\nu_1).$$

Отже, лише d_2 є баєсівським рішенням, якщо $U_x(\nu_1) < \frac{2}{3}$; лише d_1 є баєсівським рішенням, якщо $U_x(\nu_1) > \frac{2}{3}$, а у випадку $U_x(\nu_1) = \frac{2}{3}$ як d_1, так і d_2 – баєсівські рішення.

Звідси і з вигляду апостеріорних ймовірностей випливають наступні результати.

Якщо спостерігають значення $X = 1$, то для баєсівської вирішувальної функції δ^* маємо:

$$\delta^*(1) = d_2$$

за умови, що $U_1(\nu_1) < \frac{2}{3}$ чи, що те саме, за умови, що $u < \frac{8}{17}$, та

$$\delta^*(1) = d_1,$$

коли $u > \frac{8}{17}$; нарешті, для $u = \frac{8}{17}$ обидва рішення d_1 та d_2 є баєсівськими. Отже,

$$\delta^*(1) = \begin{cases} d_2, & 0 \leq u \leq \frac{8}{17}, \\ d_1, & \frac{8}{17} \leq u \leq 1. \end{cases}$$

Якщо ж спостерігають значення $X = 0$, то

$$\delta^*(0) = d_2$$

за умови $U_0(\nu_1) < \frac{2}{3}$, тобто коли $u < \frac{16}{19}$, та

$$\delta^*(0) = d_1$$

для $u > \frac{16}{19}$; нарешті, коли $u = \frac{16}{19}$, то обидва рішення d_1 та d_2 є баєсівськими. Отже,

$$\delta^*(0) = \begin{cases} d_2, & 0 \le u \le \frac{16}{19}, \\ d_1, & \frac{16}{19} \le u \le 1. \end{cases}$$

Обчислимо тепер значення баєсівського ризику $R^* = R^*(u)$ для довільної апріорної ймовірності u.

1) Якщо $0 \le u \le \frac{8}{17}$, то рішення d_2 буде баєсівським незалежно від спостережуваного значення X. Отже, згідно з таблицею для таких u маємо $R^*(u) = 5u$.

2) Якщо $\frac{8}{17} < u < \frac{16}{19}$, то $\delta^*(0) = d_2$ та $\delta^*(1) = d_1$. Тому з рівностей (9), (11) та таблиці видно, що

$$R^*(u) = U(\nu_1)R_{\nu_1}(\delta^*) + U(\nu_2)R_{\nu_2}(\delta^*)$$

$$= u\left[0 \cdot \frac{3}{4} + 5 \cdot \frac{1}{4}\right] + (1-u)\left[10 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3}\right]$$

$$= \frac{5}{4}u + \frac{10}{3}(1-u).$$

3) Якщо $\frac{16}{19} \le u \le 1$, то рішення d_1 буде баєсівським незалежно від того, яке значення X спостерігають. Згідно з таблицею, у цьому випадку $R^*(u) = 10(1-u)$.

Отже,

$$R^*(u) = \begin{cases} 5u, & 0 \le u \le \frac{8}{17}, \\ \frac{5}{4}u + \frac{10}{3}(1-u), & \frac{8}{17} \le u \le \frac{16}{19}, \\ 10(1-u), & \frac{16}{19} < u \le 1. \end{cases}$$

Ціна спостереження

У багатьох статистичних задачах прийняття рішень спостереження в.в. X пов'язано з певними витратами, які мають враховуватися статистиком при розрахунку ризику від прийняття рішень, що використовує результати спостереження X. Ця обставина грає особливо важливу роль у випадку, коли статистику треба вирішити, яку з декількох в.в. спостерігати, або вирішити, чи робити спостереження взагалі.

Нехай $c(w, x)$ позначає ціну спостереження значення x величини X, якщо $W = \nu$. Тоді, якщо u – це щільність розподілу U в.в. W, то середня ціна спостереження дорівнює

$$\mathrm{E}\left[c(W, X)\right] = \int\limits_{\Omega}\int\limits_{\mathbb{R}} c(\nu, x) f(x|\nu) u(\nu)\, dx d\nu.$$

Ми будемо припускати, що для ціни $c(W, X)$ є істинним припущення про середню корисність. Іншими словами, будемо вважати, що ця ціна виражена у відповідних одиницях від'ємної корисності так, що істотним для нас є лише середнє значення ймовірнісного розподілу ціни $c(W, X)$.

Загальним ризиком від спостереження X і прийняття вирішувальної функції δ називають суму ризику $R(U, \delta)$ і середньої ціни спостереження $\mathrm{E}[c(W, X)]$. Статистик має вибрати спостереження з деякого класу доступних спостереженню в.в. і відповідну баєсівську вирішувальну функцію δ, яка мінімізує загальний ризик.

Виражаючи загальний ризик у вигляді суми ризику вирішувальної функції δ і середньої ціни спостереження, ми неявно використовуємо припущення про адитивність корисностей статистика. Власне кажучи, всі результати в теорії статистичних рішень ґрунтуються на цьому припущенні, і ми будемо використовувати його далі.

Дуже часто статистик може вибрати той чи інший об'єм випадкової вибірки, і ціна спостереження залежить лише від цього об'єму вибірки.

Приклад

Розглянемо знову попередній приклад, і припустимо тепер, що ціна спостереження в.в. X дорівнює c, $c > 0$. Статистик може прийняти рішення, не спостерігаючи X, або заплатити суму c і спостерігати X перед прийняттям рішення. При заданому апріорному розподілі u ставиться питання, на яку суму c варто погоджуватися статистику?

Для розв'язання цієї задачі треба порівняти мінімальне значення ризику $R^*(u)$ без врахування ціни спостереження c, яке можна отримати на основі спостереження X, з мінімальним ризиком $R_0(u)$, що відповідає баєсівському рішенню за відсутності спостережень. Функція R^* вже знайдена. Функція R_0 має вигляд

$$R_0(u) = \begin{cases} 5u, & 0 \leq u \leq \frac{2}{3}, \\ 10(1-u), & \frac{2}{3} < u \leq 1. \end{cases} \quad (18)$$

$R^*(u) = R_0\,(u)$, якщо $u \leq \frac{8}{17}$ чи $u \geq \frac{16}{19}$.
Отже, при значенні апріорної ймовірності u, що лежить в одному з цих інтервалів, статистик може досягти і без спостереження X того ж значення ризику, що і при спостереженні X.

Якщо ж $\frac{8}{17} < u < \frac{16}{19}$, то $R^*(u) < R_0\,(u)$. За можливість спостереження величини X перед прийняттям рішення статистику варто погоджуватися на будь-яку ціну c таку, що $c < [R_0\,(u) - R^*\,(u)]$. Різниця між ризиками $[R_0\,(u) - R^*\,(u)]$ максимальна для $u = \frac{2}{3}$, де вона дорівнює $\frac{25}{18}$.

Приклад

Припустимо тепер, що в задачі прийняття рішень з прикладу статистик може вибирати кількість спостережень в.в. X. Іншими словами, статистик може спостерігати значення n в.в. $X_1, \dots, X_n$, причому при кожному фіксованому значенні $W = \nu_i, i = 1, 2$, величини $X_1, \dots, X_n$ н.о.р. з тим самим умовним розподілом $f(\cdot|\,\nu_i)$, що й одне спостереження X з прикладу. Умовні розподіли $f(\cdot|\,\nu_i)$ задаються рівностями (16) і можуть бути записані в такому вигляді (при $x = 0, 1$)

$$f(x|\,\nu_1) = \frac{3^x}{4}, \quad f(x|\,\nu_2) = \frac{2^{1-x}}{3}. \tag{19}$$

Для будь-якої послідовності результатів $x_1, \ldots, x_n$ спостережень $X_1, \ldots, X_n$ покладемо $y = \sum_{i=1}^{n} x_i$. Тоді значення $g\left(x_1, \ldots, x_n \middle| \nu_i\right)$ умовного розподілу в.в. $X_1, \ldots, X_n$ за умови $W = \nu_i$ такі:

$$g\left(x_1, \ldots, x_n \middle| \nu_1\right) = \frac{3^y}{4^n}, \quad g\left(x_1, \ldots, x_n \middle| \nu_2\right) = \frac{2^{n-y}}{3^n}. \tag{20}$$

Якщо ціна кожного спостереження дорівнює $c > 0$, то вибірку якого об'єму n варто обрати статистику?

Для будь-якої попередньо відомої ймовірності $u = \mathrm{P}(W = \nu_1)$ позначимо через $u(x_1, \ldots, x_n)$ апостеріорну ймовірність того, що $W = \nu_1$, за умови, що зроблено вибірку обсягу n і $X_1 = x_1, \ldots, X_n = x_n$. З теореми Баєса випливає, що

$$u(x_1, \ldots, x_n) = \frac{\frac{3^y}{4^n} u}{\frac{3^y}{4^n} u + \frac{2^{n-y}}{3^n}(1-u)} = \\ = \left[1 + \left(\frac{1-u}{u}\right)\left(\frac{8}{3}\right)^n \left(\frac{1}{6}\right)^y\right]^{-1}. \qquad (21)$$

Як було показано, при $u\left(x_1, \ldots, x_n\right) < \frac{2}{3}$ рішення d_2 є баєсівським, у протилежному випадку d_1 – баєсівське рішення. Тому з (21) видно, що баєсівська вирішувальна функція $\delta_n\left(x_1, \ldots, x_n\right)$ набуває значення d_2, якщо

$$y < \frac{\log\left[2\left(1-u\right)/u\right] + n\log\frac{8}{3}}{\log 6}. \qquad (22)$$

У протилежному випадку $\delta_n\left(x_1, \ldots, x_n\right)$ набуває значення d_1.

Нехай k_n позначає значення правої частини нерівності (22). Ризик $R(u, \delta_n)$ баєсівської вирішувальної функції δ_n задовольняє співвідношенню

$$R(u, \delta_n) = 5u\,\mathrm{P}\left(\sum_{i=1}^{n} X_i < k_n \,\middle|\, W = \nu_1\right) + \\ +10(1-u)\,\mathrm{P}\left(\sum_{i=1}^{n} X_i \geq k_n \,\middle|\, W = \nu_2\right). \qquad (23)$$

Умовний розподіл суми $\sum\limits_{i=1}^{n} X_i$ при $W = \nu_1$ є біноміальним з параметрами n і $\frac{3}{4}$. Умовний розподіл суми $\sum\limits_{i=1}^{n} X_i$ при $W = \nu_2$, також біноміальний, але з параметрами n і $\frac{1}{3}$.

Отже, ризик $R\left(u, \delta_n\right)$ можна обчислити для будь-якої попередньої ймовірності u і не занадто великого обсягу вибірки n за таблицями біноміального розподілу. Для великих значень n значення $R\left(u, \delta_n\right)$ можна обчислити, використовуючи нормальну апроксимацію біноміального закону (теорема Муавра-Лапласа) і таблиці нормального розподілу.

Для одержання загального ризику баєсівської вирішувальної функції $\delta_n(x_1, \ldots, x_n)$ при вибірці обсягу n до ризику $R(u, \delta_n)$ потрібно додати ціну вибірки nc. Оптимальний об'єм вибірки при попередній імовірності u – це значення n, яке мінімізує загальний ризик $R_t(u, \delta_n)$, що визначається формулою

$$R_t(u, \delta_n) = R(u, \delta_n) + nc. \qquad (24)$$

n	$R(u,\delta_n)$	$R_t(u,\delta_n)$	n	$R(u,\delta_n)$	$R_t(u,\delta_n)$
21	0,1394	0,3494	29	0,0609	0,3509
22	0,1370	0,3570	30	0,0513	0,3513
23	0,1116	0,3416	31	0,0506	0,3606
24	0,1056	0,3456	32	0,0411	0,3611
25	0,0903	0,3403	33	0,0402	0,3702
26	0,0822	0,3422	34	0,0334	0,3734
27	0,0738	0,3438	35	0,0313	0,3813
28	0,0646	0,3446			

У таблиці наведені деякі значення ризику $R(u,\delta_n)$ і загального ризику $R_t(u,\delta_n)$ при попередній ймовірності $u=\frac{2}{3}$, і ціні спостереження $c=0,01$. Оптимальний обсяг вибірки в нашій задачі дорівнює 25, а мінімальне значення загального ризику дорівнює 0,3403.

Спостереження в кілька етапів

Розглянемо тепер задачу, у якій спостерігаються дві випадкові величини (чи два випадкових вектори) X та Y. Нехай $f\left(x, y \middle| \nu\right)$ – це сумісний умовний розподіл ймовірностей X та Y при $W=\nu$, $\nu \in \Omega$. Апостеріорна щільність розподілу $u\left(\nu \middle| x, y\right)$ параметра W при $X=x$ та $Y=y$ обчислюється в точці ν за формулою

$$u\left(\nu \middle| x, y\right)=\frac{f\left(x, y \middle| \nu\right) u\left(\nu\right)}{\int\limits_{\Omega} f\left(x, y \middle| \nu'\right) u\left(\nu'\right) d\nu}. \tag{25}$$

Припустимо, що X та Y спостерігаються не одночасно, X спостерігається раніше від Y. Нехай $g(x|\nu)$ позначає умовний розподіл ймовірностей X при $W=\nu$. Після спостереження $X=x$ ми можемо обчислити апостеріорний розподіл $u(\nu'|x)$ для W перед спостереженням Y за формулою

$$u(\nu|x)=\frac{g(x|\nu)\,u(\nu)}{\int\limits_{\Omega} g(x|\nu')\,u(\nu')\,d\nu'}. \tag{26}$$

Далі, умовний розподіл ймовірностей $h(y|\nu,x)$ для Y при $W=\nu$ і $X=x$ має вигляд

$$h(y|\nu,x)=\frac{f(x,y|\nu)}{g(x|\nu)}. \tag{27}$$

Отже, на другому етапі експерименту, коли спостерігається Y, розподіл ймовірностей (26) можна розглядати як попередньо відомий розподіл ймовірностей для W, і умовні розподіли ймовірностей (27) утворюють при $\nu \in \Omega$ відповідне сімейство розподілів Y.

Апостеріорний розподіл ймовірностей $\xi\left(\cdot\middle| x, y\right)$ параметра W при $Y = y$ можна знайти наступним чином:

$$u\left(\nu\middle| x, y\right) = \frac{h\left(y\middle| \nu, x\right) u\left(\nu\middle| x\right)}{\int\limits_{\Omega} h\left(y\middle| \nu', x\right) u\left(\nu'\middle| x\right) d\nu'}. \tag{28}$$

Якщо розподіли ймовірностей (26) і (27) підставити в (28), то прийдемо до (25). Це означає, що якщо спостереження проводяться в декілька етапів, то апостеріорний розподіл можна обчислювати на кожному етапі, беручи як апріорний розподіл для наступного етапу апостеріорний розподіл, отриманий на попередньому етапі.

З наших міркувань також випливає, що якщо апостеріорний розподіл W при $X = x$ та $Y = y$ обчислюється в два прийоми, то остаточний результат не залежить від того, яка з випадкових величин, X чи Y, спостерігалася спочатку.

Процес прийняття рішення може бути тепер описаний у такому спрощеному вигляді. В заданий момент часу статистик має розподіли ймовірностей параметра W. З часом до статистика надходить інформація про W, і статистик використовує цю інформацію для переоцінки розподілу W. У ті моменти часу, коли статистику треба прийняти рішення, наслідки якого пов'язані з W, він вибирає рішення, оптимальне щодо розподілу W в даний момент.

Такий процес прийняття рішень досить реалістичний. Дійсно, протягом життя ми переглядаємо наші уявлення про параметри з ростом інформації про них і, приймаючи рішення, ґрунтуємося на наших сьогоднішніх уявленнях. У деяких ситуаціях, однак, від вибору рішення в даний момент може залежати та інформація, яку ми одержимо надалі, і отже, цей вибір може вплинути на рішення статистика, які він прийме в майбутньому.

Задачі, у яких статистик повинен брати до уваги майбутнє і будувати відповідні плани, називаються задачами послідовного прийняття рішень.

Вправа

Виробник спеціалізованих товарів для роздрібного ринку має вирішити, продукцію якої комплектації виготовляти у наступному році. У нього є три можливі вибори комплектації з різною вартістю виготовлення: базова, люкс та повна. Виробник обмежив накладні витрати сумою $\$1\,300\,000$.

Прибуток та витрати на виготовлення для товару кожного типу наведено в таблиці.

Комплектація	Накладні витрати	Прибуток від проданої одиниці товару
Базова	100 000	1,00
Люкс	400 000	1,20
Повна	1 000 000	1,50

Минулого року виробник продав 2 100 000 одиниць товару та підготував прогнози дохідності на наступний рік за трьома сценаріями: низької реалізації (70% минулорічного рівня продажу), середньої реалізації (такий самий рівень) та високої реалізації (на 15% більше минулорічного рівня).

i) Визначте річний прибуток для кожної можливої комбінації.
ii) Знайдіть мінімаксне рішення цієї задачі.
iii) Визначте баєсівське рішення відносно річного прибутку за умови такого ймовірнісного розподілу можливих сценаріїв: $P(\text{низька реалізація}) = 0,25$; $P(\text{середня реалізація}) = 0,6$; $P(\text{висока реалізація}) = 0,15$.

Вправа

Заявника на страхування життя згідно з анкетою класифікують як "стандартне здоров'я" (1), "погіршене здоров'я" (2) чи "не підлягає страхуванню" (3). Анкета не є ідеальним класифікатором та може віднести заявника до неправильної категорії.
Рішення записати заявника у стан i позначають через d_i, а правильний стан заявника – через θ_i.
Функція втрат для цього рішення наведена нижче.

	d_1	d_2	d_3
θ_1	0	5	8
θ_2	12	0	3
θ_3	20	15	0

i) Визначте мінімаксне рішення щодо приписання заявника до категорії.

ii) Згідно з анкетою, новий заявник належить до категорії з погіршеним здоров'ям. Але насправді, з-поміж заявників, здоров'я яких вважають погіршеним, 15% мають нормальний стан здоров'я, а 25% не підлягають страхуванню. Знайти баєсівське рішення для цього заявника.

Вправа

Страховик вивчає необхідність залучення зовнішньої компанії для проведення рекламної кампанії. Якщо він вирішить обійтися власними силами, то очікувані витрати на рекламу наступного року становитимуть $\$1\,400\,000$, внаслідок чого сподівається отримати портфель зі $100\,000$ полісів. У своєму розпорядженні страховик має розцінки двох різних компаній.

Компанія A виставила ціну $\$2\,100\,000$ на рік та вважає, що її послуги спричинять розширення бізнесу до $125\,000$ полісів.

Компанія B править $\$3\,000\,000$ на рік, вважаючи, що її послуги розширять бізнес до $140\,000$ полісів.

На поточний момент кожен поліс дає страховій компанії річний прибуток $\$30$; але ця сума не гарантується в майбутньому. Компанія оцінює, що сума залишиться на нинішньому рівні з ймовірністю $0,6$, але також може зменшитися до $\$20$ з імовірністю $0,25$ чи зрости до $\$40$ з імовірністю $0,15$.

i) Поясніть, який із трьох варіантів можна одразу відкинути.
ii) Визначте баєсівське рішення у задачі максимізації прибутку компанії в наступному році.

Вправа

Величини індивідуальних позовів за певним страховим полісом можуть набувати значення 100, 150 або 200. За рік може надійти не більше одного позову. Річна премія дорівнює 60.

Страховик повинен вибрати між трьома видами договору перестрахування:

A) не застосовувати перестрахування;

B) індивідуальний ексцедент збитку з рівнем утримання 150 та премією 10;

C) пропорційне перестрахування з рівнем утримання 25% та премією 20.

i) Складіть таблицю втрат для страховика.

ii) Визначте, чи є домінування між угодами перестрахування з точки зору страховика.

iii) Знайдіть для страховика мінімаксне рішення.

Вправа

Функція втрат у задачі прийняття рішень задана матрицею

	θ_1	θ_2	θ_3
d_1	10	15	5
d_2	8	20	15
d_3	12	15	10
d_4	5	23	8

де d_i – рішення, а θ_j – можливі стани природи.

i) Визначте рішення, яке одразу можна відкинути.

ii) Знайдіть мінімаксне рішення задачі.

iii) Визначте баєсівське рішення задачі за умови, що $P(\theta_1) = 0,4$, $P(\theta_2) = 0,25$ та $P(\theta_3) = 0,35$.

Вправа

Припустимо, що статистик має зробити висновок про те, чи деяка величина рівномірно розподілена на інтервалі $(0,1)$, чи вона рівномірно розподілена на інтервалі $\left(0,\frac{1}{2}\right)$. Ймовірність кожної з цих можливостей дорівнює $\frac{1}{2}$. Нехай збиток від прийняття правильного рішення дорівнює 0, а якщо прийняте невірне рішення, то збиток дорівнює $a>0$. Доведіть, що якщо статистик може вибирати число спостережень над випадковою величиною, причому ціна кожного спостереження дорівнює $c>0$, то йому варто робити n^* спостережень, де n^* – невід'ємне ціле число, яке мінімізує значення $a2^{-(n+1)}+nc$.

Розглянемо задачу прийняття рішень, у якій $\Omega = \{w_1, w_2\}$, $D = \{d_1, d_2, d_3\}$ а функція втрат $L(w, d)$ задається таблицею нижче. Припустимо, що нам доступне спостереження випадкової величини X з наступними умовними розподілами:

$$\mathrm{P}\left(X = 1 \middle| W = w_1\right) = \frac{3}{4}, \quad \mathrm{P}\left(X = 0 \middle| W = w_1\right) = \frac{1}{4};$$

$$\mathrm{P}\left(X = 1 \middle| W = w_2\right) = \frac{1}{4}, \quad \mathrm{P}\left(X = 0 \middle| W = w_2\right) = \frac{3}{4}.$$

Нехай $p = \mathrm{P}\left(W = w_1\right)$. Визначте для кожного значення $p\,(0 \le p \le 1)$ байєсівську вирішуючу функцію і побудуйте графік байєсівського ризику $\rho^*\left(\xi\right)$ як функції від p.

	d_1	d_2	d_3
w_1	0	10	3
w_2	10	0	3

Припустимо, що в умовах попередньої задачі статистик перед прийняттям рішення може спостерігати значення випадкових величин $X_1, \dots, X_n$, які при кожному даному значенні $W = w_i \quad (i = 1, 2)$ незалежні і однаково розподілені з X. Для довільного об'єму вибірки n знайдіть байєсівську вирішуючу функцію при кожному значенні p.

Розглянемо два ящики A і B, кожен з який містить червоні і зелені кулі. Відомо, що в одному з ящиків половина всіх куль червоні, а інші кулі зелені. В іншому ящику чверть усіх куль червоні, а три чверті куль зелені. Нехай ящик, де половина куль червоні, позначений через W, причому невідомо, $W = A$ чи $W = B$. Нехай далі, $\mathrm{P}(W = A) = p$ і $\mathrm{P}(W = B) = 1 - p$, де p – задане число, $0 < p < 1$. Припустимо, що статистик може вибрати навмання одну кулю з ящика A чи B і після цього спостереження повинний прийняти рішення $W = A$ чи $W = B$. Покажіть, що якщо $\frac{1}{2} < p < \frac{2}{3}$, то для того, щоб максимізувати ймовірність правильного рішення, йому варто вийняти кулю з ящика B. Покажіть, далі, що якщо $\frac{2}{3} \leq p \leq 1$, то вибір ящика несуттєвий.