Спецкурс "Теория риска" (для 409 гр.)

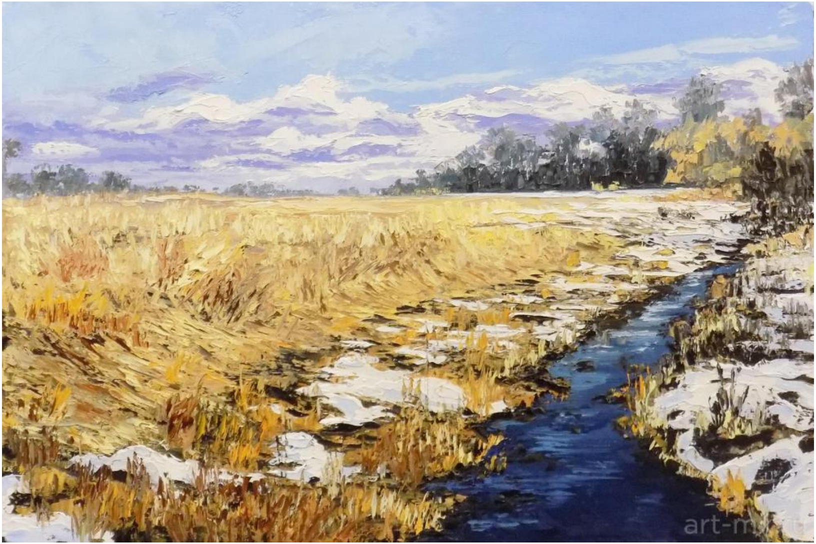
Проф. Екатерина Вадимовна Булинская

(МГУ имени М.В.Ломоносова)

Лекция 11

Москва, 25 ноября 2020 г.

С ДОБРЫМ УТРОМ!



План лекции

- Experience rating (апостериорная тарификация)
- Американская достоверность или теория ограниченных флуктуаций
- Современная или европейская достоверность
- Модель Бюлмана с одним контрактом

Апостериорная тарификация

Рассмотренные ранее тарифные принципы были разработаны в предположении, что портфель страховой компании однородный, т.е. отдельные риски имеют одинаковое распределение.

Однако на практике это далеко не так. Например, в автомобильном страховании число происшествий у отдельных водителей зависит от многих факторов, таких как возраст, место жительства, цели использования автомобиля, частота поездок и др. Поэтому во многих странах используется так называемая система бонус-малус, когда при отсутствии происшествий за год понижается тариф в следующем году (бонус), а за наличие происшествий тариф повышается (малус).

Это как раз пример "апостериорной" тарификации, основанной на результатах наблюдений за числом страховых случаев по данному контракту в предыдущие годы (experience rating).

Возникает вопрос, как оптимальным образом присоединить дополнительную информацию, касающуюся прошлых происшествий, к априорным сведениям о риске, которые уже были включены в тарификацию, называемую "априорной" , и с их помощью контракт был отнесен к определенной тарифной категории.

Параметры риска (например, чистая премия) могут быть определены тем более точно и надежно, чем дольше период наблюдения за контрактом (при этом нужна однородность по времени) и/или чем больше контрактов наблюдается (а это означает однородность портфеля "в пространстве").

В большинстве случаев страхования не жизни нет ни той, ни другой однородности. Риски изменяются со временем, и даже внутри одного тарифного класса нет полной однородности. Таким образом, надо уметь извлекать максимум информации о риске из имеющихся данных, комбинируя оба измерения, временное и пространственное.

При тарификации можно предусмотреть простой подсчет отношения общего размера возмещений по страховым случаям к числу единиц риска (более общим образом, объему риска). Но на практике такой элементарный метод оказывается неуместным, когда происходит детализация тарифной системы и сталкиваются с подклассом, содержащим недостаточное число (или объем) единиц риска для того, чтобы можно было пренебречь случайными колебаниями.

Теория достоверности (credibility)

Mowbray A.H. (1914) – how many trials/results need to be observed before I can believe my data?

Whitney A.W. (1918) – focus was on combining existing estimates and new data to derive new estimates.

Итак, возникает задача установить, как распределена величина возмещений для неоднородного портфеля. Цель теории достоверности (credibility theory) состоит в том, чтобы найти решение этой главной задачи в страховании не жизни и предоставить теоретическое обоснование его справедливости.

Основные идеи предлагаемого подхода не являются новыми. Еще в 1910-1920гг. заокеанские актуарии разработали технику тарификации, которую назвали теорией ограниченных флуктуаций.

Она состояла в том, что тариф контракта, входящего в некоторую совокупность, составляли путем взвешивания, т.е. выпуклой комбинации данных по всей совокупности и относящихся непосредственно к данному контракту.

А именно, тариф имел следующий вид:

$$(1-\zeta)m + \zeta \bar{X}_t,$$

где m - чистая премия, определяемая для всей совокупности, а $\bar{X}_t$ - это средний размер возмещений по рассматриваемому контракту, полученный по наблюдениям в течение t лет. Коэффициент ζ, который еще с тех времен называется "коэффициентом достоверности" , это число, заключенное между 0 и 1, при этом он тем ближе к 1, чем "достовернее" данные, касающиеся непосредственно изучаемого контракта.

Интересный частный случай - это тот, для которого $\zeta = 1$ (т.е. случай полной достоверности - full credibility).

Это решение, остроумное и полезное для практики, обладало одним важным недостатком: оно было эмпирическим. Отсутствие теоретического обоснования приводило к тому, что было неизвестно, при каких предположениях (гипотезах) это решение справедливо.

Только в 60-е годы прошлого века Ханс Бюлман (Hans Bühlman) заложил прочные математические основы теории достоверности, выделив необходимые для нее гипотезы. Эта теория называется с тех пор современной теорией достоверности или "европейской достоверностью."

В последующие годы была получена масса значительных результатов, позволивших расширить область применимости теории и учитывать все более разнообразные ситуации, возникающие на практике. Таким образом, произошел переход от эмпирической техники к настоящей математической теории. Отметим, однако, что возможности практического применения ряда моделей ограничивается тем, что качество используемых оценок оставляет желать лучшего.

Закончим это краткое введение рассмотрением примера, предложенного Норбергом (R.Norberg).

Пример Норберга

Пусть имеется множество из 20 независимых идентичных контрактов (т.е. априори нет никаких оснований их различать). Страховые возмещения, которые возможно придется выплатить, равны 1 для любого контракта и любого происшествия. Число происшествий за год по любому контракту - бернуллиевская случайная величина, вероятность наступления происшествия равна θ, а вероятность его отсутствия $1-\theta$.
Наблюдения в течение 10 лет дали следующие результаты по числу происшествий над контрактами:

Всего: 0 0 2 0 2 0 2 0 6 1 4 3 1 1 0 0 5 1 1 0

Общее число происшествий (= сумме возмещений) равно 29. В результате средняя чистая премия за год для одного контракта оказывается 29/200=0,145.

Заметим, что если рассмотреть портфель из 17 контрактов, исключив 9, 11 и 17 контракты, по которым больше всего происшествий, то чистая премия оставшихся будет равна 14/170=0,083, что значительно меньше получившейся ранее.

Возникают 2 вопроса:

1. Портфель из 20 контрактов, который априори представлялся однородным, является ли таким, в самом деле, апостериори?

Обычные статистические процедуры позволяют дать ответ на этот вопрос.

Пусть θ_j - параметр бернуллиевской величины, описывающей происшествия j-го контракта.

$$\hat{\theta}_j = \frac{\text{число происшествий по } j - \text{му} \quad \text{контракту}}{10}$$

- несмещенная оценка с минимальной дисперсией для параметра θ_j.

$$\hat{\theta} = \frac{1}{20}\sum_{j=1}^{20}\hat{\theta}_j = 0,145.$$

Необходимо проверить гипотезу

$$H_0 : \theta_1 = \theta_2 = \ldots = \theta_{20}.$$

При этой гипотезе статистика

$$X^2 = 10 \sum_{j=1}^{20} (\hat{\theta}_j - \hat{\theta})^2 / \hat{\theta}(1 - \hat{\theta})$$

имеет хи-квадрат распределение с 19 степенями свободы.

В данном случае $X^2 = 49,16$, а 5%-я доверительная граница $\chi^2_{19} = 30,1$. Следовательно, гипотезу H_0 приходится отбросить.

2. Как же выбирать тариф для портфеля, который априори представлялся однородным, но оказался неоднородным апостериори?

Введение одинакового тарифа для всех (т.е. весь суммарный ущерб равномерно распределяется по контрактам), подсказываемое априорной однородностью портфеля приводит к чистой премии $\hat{\theta} = 0,145$, не являющейся справедливой. В самом деле, по контрактам 9, 11, 17 тариф оказывается сильно заниженным, в то время как по остальным завышенным. А это может привести к тому, что кто-то из владельцев этих контрактов отправится в другую компанию на поиски более выгодных условий, что еще более усилит неоднородность портфеля.

Тарификация каждого контракта лишь в зависимости от числа происшествий по нему (т.е. когда заставляют каждого застрахованного платить сумму ущерба от его происшествий) может привести к тому, что чистая премия будет очень сильно меняться, особенно по контрактам с малым числом происшествий.

Возможны 2 решения:

- взять $\hat{\theta}$ и $\hat{\theta}_j$ с такими весами, которые позволяют учесть оба типа информации, а результат не будет подвержен сильной изменчивости,

- выделить более однородные части портфеля (подпортфели), а затем осуществлять взвешивание информации.

Теория ограниченных флуктуаций

Разработанная еще в 1910-1920гг. и используемая с тех пор американскими актуариями техника достаточно проста.

Начнем с рассмотренного примера. Тариф контракта с номером j задается выпуклой комбинацией

$$(1-\zeta_t)\hat{\theta}+\zeta_t\cdot{}_j\bar{N}_t.$$

В этой формуле $\hat{\theta}$ - это оценка чистой премии по всему портфелю, а ${}_j\bar{N}_t=\frac{1}{t}\sum_{j=1}^{t}{}_jN_t$ - это среднее число происшествий за год, связанное с контрактом j, по наблюдениям за t лет.

Возникает вопрос, наблюдениями за сколько лет надо обладать, для того чтобы относительная величина случайных флуктуаций ${}_j\bar{N}_t$ была ограниченной.

Пусть заданы $\alpha > 0$ и $\varepsilon > 0$ (например, $\alpha = 10^{-1}$ и $\varepsilon = 10^{-1}$), требуется определить такое t, чтобы

$$P\left(\left|\frac{\theta_j - {}_j\bar{N}_t}{\theta_j}\right| \leq \alpha\right) \geq 1 - \varepsilon.$$

Если предположить, что применима центральная предельная теорема, т.е. что можно считать нормированную величину

$$\frac{\sqrt{t} \cdot ({}_j\bar{N}_t - \theta_j)}{\sqrt{\theta_j(1-\theta_j)}}$$

распределенной по закону $\mathcal{N}(0,1)$, решение легко найти

$$t \geq t_j(\varepsilon, \alpha) = \left(\frac{\Phi^{-1}(1-\varepsilon/2)}{\alpha}\right)^2 \cdot \frac{1-\theta_j}{\theta_j}.$$

Заметим, что t_j пропорционально коэффициенту вариации параметра θ_j. Таким образом, чем больше разброс значений параметра, тем большее число лет t_j надо наблюдать.

$\alpha \diagdown \theta_j$	0,1	0,2	0,5
0,1	2248	1088	272
0,2	613	272	68
0,5	98	44	11

Значения $t_j(\varepsilon, \alpha)$ для $\varepsilon = 0,1$.

Итак, если наблюдения длились не менее t_j лет, то можно быть уверенным (с большой вероятностью), что не будет больших случайных колебаний. Следовательно, при $t \geq t_j$ можно взять $\zeta_t = 1$, т.е. мы приходим к случаю полной достоверности (full credibility).

В случае $t < t_j$ также ставится цель ограничить относительную ошибку слагаемого, связанного с наблюдениями за числом происшествий, т.е. ищется наибольшее ζ_t такое, что

$$P\left(\zeta_t \left|\frac{\theta_j - {}_j\bar{N}_t}{\theta_j}\right| \le \alpha\right) \ge 1 - \varepsilon.$$

Проводя вычисления, аналогичные предыдущим, получаем

$$\zeta_t^2 \le \frac{t}{t_j(\varepsilon, \alpha)}.$$

Отсюда наибольшее значение оказывается $\zeta_t = \sqrt{t/t_j}$. Следовательно, в любом случае

$$\zeta_t = \min\left(\sqrt{\frac{t}{t_j(\varepsilon, \alpha)}}, 1\right).$$

Современная теория достоверности

Ханс Бюлман в 1967г. заложил прочный математический фундамент теории достоверности, использовав совсем иной подход.

Пусть рассматривается страховой контракт, принадлежащий однородному портфелю. Характеристики риска, соответствующего контракту, зависят от случайного параметра Θ. Распределение U этого параметра называется функцией структуры портфеля.

Каждый год наблюдается число происшествий, связанных с контрактом, т.е. реализация случайной величины X_i, $1 \leq i \leq t$.

Гипотезы (B)

Случайные величины $X_1, \ldots, X_t$ имеют конечный второй момент ($\in L^2$). Условно (относительно Θ) они независимы и одинаково распределены. Обозначения:

$$\mu(\Theta) = \mathsf{E}(X_i|\Theta),$$
$$m = \mathsf{E}(\mu(\Theta)) = \mathsf{E}X_i,$$
$$a = \mathsf{D}(\mathsf{E}(X_i|\Theta)) = \mathsf{D}(\mu(\Theta)),$$
$$\sigma^2(\theta) = \mathsf{D}(X_i|\Theta = \theta),$$
$$s^2 = \mathsf{E}(\mathsf{D}(X_i|\Theta)) = \mathsf{E}(\sigma^2(\Theta))$$

Можно заметить, что a измеряет неоднородность, обусловленную разбросом функции структуры U. Так, в случае $dU(\theta) = \delta_{\theta_0}$, параметр $a = 0$. Напротив, s^2 служит для измерения (средней) неоднородности данных.

Отметим также, что, в отличие от гипотетического априорного распределения в байесовском подходе в статистике, функция структуры U действительно существует, хотя она может быть неизвестна. Причем нам придется находить явный вид оценок некоторых ее параметров, например,

$$a = \int (\mu(\theta) - m)^2 \, dU(\theta).$$

Задача, которую надо решить.
На основе реализации вектора $(X_1, \ldots, X_t)$, т.е. наблюдения за числом происшествий по контракту в течение t лет,
а) определить чистую премию $\mu(\theta)$ контракта,
б) предсказать ущерб X_{t+1} в следующем году.

Оптимальный прогноз

Теорема

$Z^* = E(X \mid \mathcal{A})$ - наилучший в среднем квадратичном прогноз случайной величины X с помощью $\mathcal{A}$-измеримых величин Z.

Док-во. По определению $E(X - Z^*)^2 = \inf_Z E(X - Z)^2$. Поскольку $E(X \pm E(X|\mathcal{A}) - Z)^2 = E(X - E(X|\mathcal{A}))^2 + E(E(X|\mathcal{A}) - Z)^2 + 2E(X - E(X|\mathcal{A}))(E(X|\mathcal{A}) - Z)$, где последнее слагаемое равно нулю, а первое не включает Z, то минимум достигается при $Z = Z^*$. $\square$

Согласно доказанной теореме решение сформулированной задачи - это $\mathrm{E}(\mu(\Theta)|X_1, \ldots, X_t)$ и $\mathrm{E}(X_{t+1}|X_1, \ldots, X_t)$. Однако в большинстве случаев неизвестно, как их сосчитать.

Частный случай

Если при фиксированном $\Theta = \theta$ существует условная плотность $f(x|\theta)$ распределения X_i, то формула Байеса дает

$$E(\mu(\Theta)|X_1 = x_1, \ldots, X_t = x_t) = \frac{\int \mu(\theta)(\Pi_{i=1}^t f(x_i|\theta))\, dU(\theta)}{\int (\Pi_{i=1}^t f(x_i|\theta))\, dU(\theta)}$$

Но в общем случае эта формула не может быть использована, так как надо знать f и U.

Точная достоверность (Jewell 1974)

Предположим, что закон распределения X_i при условии $\Theta = \theta$ принадлежит экспоненциальному семейству, т.е. $f(x|\theta) = c(\theta)e^{\theta x}h(x)$ - плотность по некоторой σ-конечной мере λ,

а $u_{a,b}(\theta) = K(a,b)c(\theta)^b e^{a\theta}$, $a > 0$, $b > 0$, $\theta \in R^+$ - плотность U по мере Лебега (сопряженная в байесовском смысле), $u_{a,b}(0) = u_{a,b}(\infty) = 0$.

Теорема

При указанных предположениях

1) Закон распределения Θ *при условии* $X_i = x_i$, $i = \overline{1,t}$, *- это закон* U *с параметрами* $a + \sum_{i=1}^{t} x_i$, $b + t$.

2) Отсюда

$$E(\mu(\Theta)|X_1 = x_1, \ldots, X_t = x_t) = (1 - \zeta_t)m + \zeta_t \overline{X_t},$$

где $\overline{X_t}$ *- среднее арифметическое наблюдений* $X_1, \ldots, X_t$, $m = a/b$, $\zeta_t = t/(b + t)$.

Доказательство

1) Поскольку
$$\int_{R^+} f(x|\theta)\, d\lambda(x) = 1,$$
используя явный вид плотности, запишем
$$\frac{1}{c(\theta)} = \int_{R^+} e^{x\theta} h(x)\, d\lambda(x),$$
откуда
$$-\frac{c'(\theta)}{c(\theta)^2} = \int_{R^+} x e^{x\theta} h(x)\, d\lambda(x) = \frac{\mu(\theta)}{c(\theta)},$$
следовательно,
$$\mu(\theta) = -\frac{c'(\theta)}{c(\theta)}.$$

С другой стороны, дифференцируя структурную плотность, имеем
$$u'_{a,b}(\theta) = K(a,b)e^{a\theta}(bc'(\theta)c(\theta)^{b-1} + ac(\theta)^b)$$
$$= u_{a,b}(\theta)(bc'(\theta)/c(\theta) + a) = u_{a,b}(\theta)(-b\mu(\theta) + a)$$

$$\int_0^{+\infty} u'_{a,b}(\theta)\,d\theta = u_{a,b}(+\infty) - u_{a,b}(0) = 0,$$

Следовательно,

$$0 = -b\int_0^{+\infty} \mu(\theta) u_{a,b}(\theta)\,d\theta + a.$$

Окончательно, $E(X_1) = E(\mu(\Theta)) = m = a/b$.

По формуле Байеса условное распределение Θ при $X_1 = x_1, \ldots, X_t = x_t$ равно

$$\frac{\prod_{i=1}^t f(x_i|\theta) u_{a,b}(\theta)}{\int_0^{+\infty} \prod_{i=1}^t f(x_i|\theta) u_{a,b}(\theta)\,d\theta}$$

$$= K_t(a, b, x_1, \ldots, x_t) e^{\theta(a+\sum_{i=1}^t x_i)} c(\theta)^{b+t},$$

т.е. первая часть теоремы доказана.

2) Поскольку получилось распределение такое же, как у Θ, но с новыми параметрами, то

$$E(\mu(\Theta)|X_1 = x_1, \ldots, X_t = x_t) = \frac{a+\sum_{i=1}^t x_i}{b+t}$$

$$= \frac{a}{b} \cdot \frac{b}{b+t} + \frac{t}{b+t} \cdot \frac{1}{t} \sum_{i=1}^t x_i. \ \square$$

Замечание. Удалось не только найти у.м.о. (наилучшую оценку), но она оказалась (неоднородной) линейной комбинацией результатов наблюдений x_i, $i = \overline{1, t}$. Именно поэтому, как скоро будет ясно, данный случай называют точной достоверностью.

Пуассоновское распределение со случайным параметром

Пусть мера λ сосредоточена в целочисленных точках, тогда закон Пуассона с параметром θ имеет вид

$$e^{-\theta}\frac{\theta^x}{x!}d\lambda(x) = e^{-\theta}e^{x\ln\theta}\frac{1}{x!}d\lambda(x)$$

Замена $\theta' = \ln\theta$, т.е. $\theta = e^{\theta'}$ и $d\theta' = d\theta/\theta$, приводит к записи сопряженного распределения

$$K(a,b)e^{-be^{\theta'}}e^{a\theta'}d\theta' = K(a,b)e^{-b\theta}e^{a\ln\theta}\frac{d\theta}{\theta}\mathbb{I}_{]0,+\infty[}(\theta)$$
$$= K(a,b)e^{-b\theta}\theta^{a-1}\mathbb{I}_{]0,+\infty[}(\theta)d\theta,$$

т.е. получилось гамма-распределение с параметрами (a,b). Значит, применима формула точной достоверности.

Общий случай

Поскольку неизвестно, как подсчитать $\mathsf{E}(\mu(\Theta|X_1, \ldots, X_t)$ (наилучший прогноз), т.е. проекцию в гильбертовом пространстве L^2 случайной величины $\mu(\Theta)$ на подпространство, порожденное случайными величинами $g(X_1, \ldots, X_t)$, где g пробегает все множество борелевских функций, отображающих $\mathbb{R}_t$ в $\mathbb{R}$, неизвестно, было предложено ограничиться более узким классом случайных величин:

- неоднородные линейные комбинации наблюдений $X_1, \ldots, X_t$ (векторное подпространство, порожденное $1, X_1, \ldots, X_t$),
- однородные линейные комбинации наблюдений $X_1, \ldots, X_t$ (векторное подпространство, порожденное $X_1, \ldots, X_t$).

H.Bühlmann, 1967, линейная аппроксимация

Был установлен следующий результат

Теорема

При выполнении (B) наилучшее в среднем квадратичном линейное (неоднородное) приближение

$$\widehat{\mu(\Theta)}_t^{(a)} = (1-\zeta_t)m + \zeta_t \overline{X}_t,$$

где

$$\overline{X}_t = \frac{1}{t}\sum_{i=1}^{t} X, \quad \zeta_t = \frac{at}{s^2+at} = \frac{cov(X_1, X_2)}{D(\overline{X}_t)},$$

ошибка аппроксимации

$$||\mu(\Theta) - \widehat{\mu(\Theta)}_t^{(a)}||_{L^2}^2 = D(\mu(\Theta) - \widehat{\mu(\Theta)}_t^{(a)}) = (1-\zeta_t)a.$$

Вспомогательный результат

Лемма

1. $cov(X_i, X_j) = \delta_{ij}s^2 + a,\ (\delta_{ij} = 1,\ i = j,\ \delta_{ij} = 0,\ i \neq j),$
2. $cov(\mu(\Theta), X_i) = a,$
3. $cov(\overline{X_t}, X_i) = (s^2/t) + a,\ \overline{X_t} = \frac{1}{t}\sum_{i=1}^{t} x_i,$
4. $D(\overline{X_t}) = (s^2/t) + a.$

Док-во. 1) $cov(X_i, X_j) = E(X_i - EX_i)(X_j - EX_j)$
$= E(X_i \pm E(X_i|\Theta) - EX_i)(X_j \pm E(X_j|\Theta) - EX_j)$
$= Ecov(X_i, X_j|\Theta) + cov(E(X_i|\Theta), E(X_j|\Theta))$
(остальные 2 слагаемые равны 0)
$= \delta_{ij}E\sigma^2(\Theta) + D\mu(\Theta) = \delta_{ij}s^2 + a.$

Использованы свойства у.м.о.: $EX = EE(X|\Theta)$, а также $E(XY|\Theta) = YE(X|\Theta)$, если Y измерима относительно сигма-алгебры, порожденной Θ.

2) Аналогично $cov(\mu(\Theta), X_i) = E(\mu(\Theta) - m)(X_i - m) =$
$EE[(\mu(\Theta) - m)(X_i - m)|\Theta] = E[(\mu(\Theta) - m)E((X_i - m)|\Theta)]$
$= E[(\mu(\Theta) - m)^2] = D\mu(\Theta) = a.$

3) $cov(\overline{X_t}, X_i) = (1/t)\sum_{j=1}^t cov(X_j, X_i)$
$= (1/t)\sum_{j=1}^t (a + s^2\delta_{ij}) = a + (s^2/t).$

4) $D(\overline{X_t}) = cov(\overline{X_t}, \overline{X_t}) = (1/t)\sum_{i=1}^t cov(\overline{X_t}, X_i)$
$= (1/t)\sum_{i=1}^t [a + (s^2/t)] = a + (s^2/t).$ $\square$

Доказательство теоремы Бюлмана

Пусть V - подпространство гильбертова пространства L^2, порожденное $\{1, X_1, \ldots, X_t\}$.

Проекция $\mu(\Theta)$ на V записывается в виде

$$Z = c_0 + \sum_{j=1}^{t} c_j X_j,$$

где $c_i \in R$ и удовлетворяют условиям ортогональности

$$E((\mu(\Theta) - Z)X_i) = 0$$

для $i = \overline{0, t}$ (положено $X_0 = 1$).

Для $i = 0$ получаем $c_0 + \sum_{j=1}^{t} c_j m = m$, т.е. с.в. $\mu(\Theta) - Z$ центрирована, а $c_0 = m(1 - \sum_{j=1}^{t} c_j)$.

Для $i \geq 1$ $cov(\mu(\Theta) - \sum_{j=1}^{t} c_j X_j, X_i) = 0$, иначе

$$cov(\mu(\Theta), X_i) - \sum_{j=1}^{t} c_j cov(X_j, X_i) = 0$$

$$a - \sum_{j=1}^{t} c_j(a + \delta_{ij} s^2) = 0,$$

$c_i s^2 = a(1 - \sum_{j=1}^{t} c_j)$, т.е. $c_1 = c_2 = \ldots = c_t$.

Окончательно,

$$c_i = \frac{a}{s^2 + at}, \ i \geq 1, \quad c_0 = \left(1 - \frac{at}{s^2 + at}\right) m,$$

$$\widehat{\mu(\Theta)}_t^{(a)} = \left(1 - \frac{at}{s^2 + at}\right) m + \frac{at}{s^2 + at} \cdot \frac{1}{t} \sum_{j=1}^{t} X_j,$$

т.е. получен требуемый результат.

В силу центрированности с.в. $\mu(\Theta) - \widehat{\mu(\Theta)}_t^{(a)}$ ошибка аппроксимации

$$e^2(t) = ||\mu(\Theta) - \widehat{\mu(\Theta)}_t^{(a)}||_{L^2}^2 = \mathrm{D}(\mu(\Theta) - \widehat{\mu(\Theta)}_t^{(a)}).$$

Используя свойства дисперсии и вспомогательную лемму, получим

$$e^2(t) = D(\mu(\Theta) - \zeta_t \overline{X_t})$$
$$= D(\mu(\Theta)) + \zeta_t^2 D(\overline{X_t}) - 2\zeta_t cov(\mu(\Theta), \overline{X_t})$$
$$= a + \zeta_t^2 (a + (s^2/t)) - 2a\zeta_t.$$

В силу того, что

$$\zeta_t = \frac{at}{s^2 + at}$$

получаем требуемый результат

$$e^2(t) = a + a\zeta_t - 2a\zeta_t = (1 - \zeta_t)a = \frac{as^2}{s^2 + at}. \square$$

Частные случаи

- Если $a = 0$, $s^2 > 0$, то $\mu(\Theta) = m$ п.н. Нет никакой неопределенности относительно $\mu(\Theta)$. А тот факт, что $s^2 > 0$ свидетельствует о наличии флуктуаций в наблюдениях за происшествиями. Очевидно, что m - наилучшее приближение для что $\mu(\Theta)$. Тот же ответ дает теорема, так как в этом случае $\zeta_t = 0$.
- Если $a \to \infty$, $s^2 > 0$ (на практике a очень велико по сравнению с s^2), то имеется значительный разброс $\mu(\Theta)$ вокруг среднего (т.е. априорный тарифный класс достаточно неоднородный). Значит, единственное значение m не является "хорошей оценкой"для $\mu(\Theta)$. Поскольку изменчивость наблюдений гораздо меньше, то предпочтительно оценить с помощью $\overline{X_t}$. Именно это дает теорема, так как $\zeta_t \to 1$ при $a \to \infty$.
- Если $a > 0$, $s^2 = 0$, то существует разброс $\mu(\Theta)$ вокруг m, в то время как $D(X_j|\Theta) = 0$, т.е. $X_j = \mu(\Theta)$ п.н. Естественно приблизить с помощью $\overline{X_t} = \mu(\Theta)$, а это и дает теорема, так как $\zeta_t = 1$.

- Если $a > 0$, $s^2 \to \infty$ (на практике s^2 очень велика по сравнению с a). Это означает, что разброс наблюдений гораздо более значителен, чем отклонения $\mu(\Theta)$ от m. За неимением лучшего предпочитается приблизить $\mu(\Theta)$ с помощью m. Действительно, в данном случае $\zeta_t \to 0$ при $s^2 \to \infty$, но ошибка аппроксимации не стремится к нулю. Совершенно естественно, $e^2(t) \to a$.
- Если $a = 0$, $s^2 = 0$, то $D(X_j|\Theta) = 0$, т.е. $X_j = \mu(\Theta)$ п.н. Поскольку обе величины m и $\overline{X_t}$ равны $\mu(\Theta)$ п.н. не важно, какую выпуклую комбинацию использовать, так как $(1-\zeta_t)m + \zeta_t\overline{X_t} = (1-\zeta_t)m + \zeta_t m = m = \mu(\Theta)$. Впрочем, коэффициент достоверности ζ_t не определен в данной ситуации.

Асимптотика $\widehat{\mu(\Theta)}_t^{(a)}$ при $t \to \infty$

Теорема

При гипотезе (B) справедливы утверждения:

1. *Для* $\{X_n\}_{n\geq 1}$ *выполнен УЗБЧ*

   $\lim_{t\to\infty} \overline{X_t} = \mu(\Theta)$ *п.н. и в ср.кв.*

2. $\lim_{t\to\infty} \widehat{\mu(\Theta)}_t^{(a)} = \mu(\Theta)$ *п.н. и в ср.кв.*

Доказательство

1) Случайная величина

$$\overline{X_t} - \mu(\Theta) = \frac{1}{t}\sum_{i=1}^{t}(X_i - \mu(\Theta))$$

центрирована, поэтому

$$\mathsf{E}(\overline{X_t} - \mu(\Theta))^2 = \mathsf{D}(\overline{X_t} - \mu(\Theta)) = \mathsf{D}\overline{X_t} + \mathsf{D}\mu(\Theta) - 2cov(\overline{X_t}, \mu(\Theta)).$$

Применяя вспомогательную лемму, получаем

$$\mathsf{E}(\overline{X_t} - \mu(\Theta))^2 = a + \frac{s^2}{t} + a - 2a = \frac{s^2}{t} \to 0,$$

когда $t \to \infty$, тем самым установлена сходимость $(\overline{X_t})_{t \geq 1}$ в среднем квадратичном к $\mu(\Theta)$.

Обратимся теперь к сходимости почти наверное.

Для $t \in \mathbb{N}$ обозначим σ-алгебру, порожденную случайными величинами $\{X_{\sigma(1)}, \ldots, X_{\sigma(t)}; \sigma \in \sum_t\}$, где $\sum_t$ - группа перестановок $\{1, 2, \ldots, t\}$, через φ_t.

$$\mathsf{E}(X_1|\varphi_t) = \mathsf{E}\left(\frac{X_1 + \ldots + X_t}{t}|\varphi_t\right) = \frac{X_1 + \ldots + X_t}{t} \text{ п.н.}$$

Последовательность $\{\varphi_t\}_{t\geq}$ возрастает к φ_∞ и $\lim_{t\to\infty} \mathsf{E}(X_1|\varphi_t) = \mathsf{E}(X_1|\varphi_\infty)$ почти наверное.

Таким образом, $\lim_{t\to\infty} \overline{X_t} = \mathsf{E}(X_1|\varphi_\infty)$ почти наверное. Поскольку мы установили, что предел последовательности $(\overline{X_t})_{t\geq 1}$ в среднем квадратичном равен $\mu(\Theta)$, то

$$\lim_{t\to\infty} \overline{X_t} = \mu(\Theta) = \mathsf{E}(X_1|\varphi_\infty)$$

почти наверное.

2) Ошибка аппроксимации $e(t)$ имеет вид

$$e(t) = ||\mu(\theta) - \widehat{\mu(\Theta)}_t^{(a)}||_{L^2} = \sqrt{(1-\zeta_t)a}.$$

- Если $a = 0$, то $e(t) = 0$ и $\widehat{\mu(\Theta)}_t^{(a)} = \mu(\Theta) = m$ п.н. для любого t.

- Если $a > 0$, то $\lim_{t\to\infty} \zeta_t = 1$, откуда $\lim_{t\to\infty} e(t) = 0$. Значит, $\lim_{t\to\infty} \widehat{\mu(\Theta)}_t^{(a)} = \mu(\Theta)$ в среднем квадратичном.

Вернемся к равенству

$$\widehat{\mu(\Theta)}^{(a)} = (1-\zeta_t)m + \zeta_t \overline{X_t}.$$

Поскольку $\zeta_t \to 1$ и $\overline{X_t} \to \mu(\Theta)$, когда $t \to \infty$, отсюда вытекает требуемый результат. □

Замечания

- В классической теории риска $\mu(\Theta)$ - это чистая премия (премия риска) на покрытие расходов по контракту. Оценка $\widehat{\mu(\Theta)}_t^{(a)}$ обладает 2 свойствами:

$$E\widehat{\mu(\Theta)}_t^{(a)} = E\mu(\Theta) = m$$

$$\lim_{t\to\infty} \widehat{\mu(\Theta)}_t^{(a)} = \mu(\Theta) \text{ п.н.}$$

Иначе говоря, в среднем тот же размер выплат, что при известных функции структуры и $\mu(\Theta)$, а в пределе получается премия риска.

- Поскольку при $1 \leq j \leq t$,
  $cov(X_{t+1}, X_j) = a = cov(\mu(\Theta), X_j)$,
  аналогичное док-во позволяет установить, что
  $\widehat{X_{t+1}} = (1 - \zeta_t)m + \zeta_t \overline{X}_t$,
  т.е. иметь прогноз будущих затрат.