



Лекция 3

Количественные методы
прогнозирования показателей
в бизнесе («статистическое
прогнозирование»)

Курс лекций:

«Прогнозирование и моделирование в бизнесе»

(с использованием Microsoft Excel)

Подготовил:
к.э.н. Литвин Ю.В.

Понятие рентабельности и доходности (1) (определение)

Рентабельность (ROS) – величина, рассчитываемая как частное от деления прибыли (NP) на выручку (Re):

$$ROS = \frac{NP}{Re} * 100\%$$

Доходность (Profity) - величина, характеризующая отношение прибыли (NP) к осуществленным вложениям (I):

$$Profity = \frac{NP}{I} * 100\%$$

Примеры?

Исследование и прогнозирование числовых данных (1)

Пример: рентабельность компаний цементного производства

Год	Компания А	Компания Б
2009	7,4 %	6,2 %
2010	8,2 %	9,7 %
2011	- 4,3 %	- 5,8 %
2012	2,4 %	3,5 %
2013	18,3 %	20,6 %
2014	7,3 %	4,5 %
2015	8,8 %	9,7 %
2016	9,1 %	9,7 %
2017-...	???	???

Вопросы:

1. Изменяется ли рентабельность компаний с течением времени?
2. Каковы наибольшее и наименьшее значения рентабельности? В какие годы функционирования компаний они были достигнуты?
3. Существуют ли какие-либо «выбросы» в имеющихся совокупностях данных? Объясните, с чем они могут быть связаны?
4. Какая из компаний на Ваш взгляд более привлекательна для инвестора?
5. Каковы ожидаемые значения рентабельности в 2017г. (при сохранении текущих условий)?
6. В каких диапазонах лежат ожидаемые рентабельности компаний в случае сохранения текущих условий?

Исследование числовых данных

В большинстве случаев данные концентрируются около некоторой центральной точки. Чтобы охарактеризовать набор данных необходимо выбрать какую-то величину, которая давала бы общее представление о распределении частот. Эту величину называют **средним значением**.

Рассмотрим четыре числовые характеристики, используемые для оценки среднего значения:

- выборочное среднее;
- медиана;
- мода;
- скользящее среднее.

Выборочное среднее (среднее арифметическое)

Пусть выборка состоит из чисел: X_1, X_2, \dots, X_n

Выборочное среднее (\bar{X}) – деление суммы наблюдаемых числовых величин на их количество:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \sum_{i=1}^n X_i / n$$

где \bar{X} – выборочное среднее,
 n – объем выборки,
 X_i – i -й элемент выборки,
 $\sum_{i=1}^n X_i$ – сумма всех элементов выборки

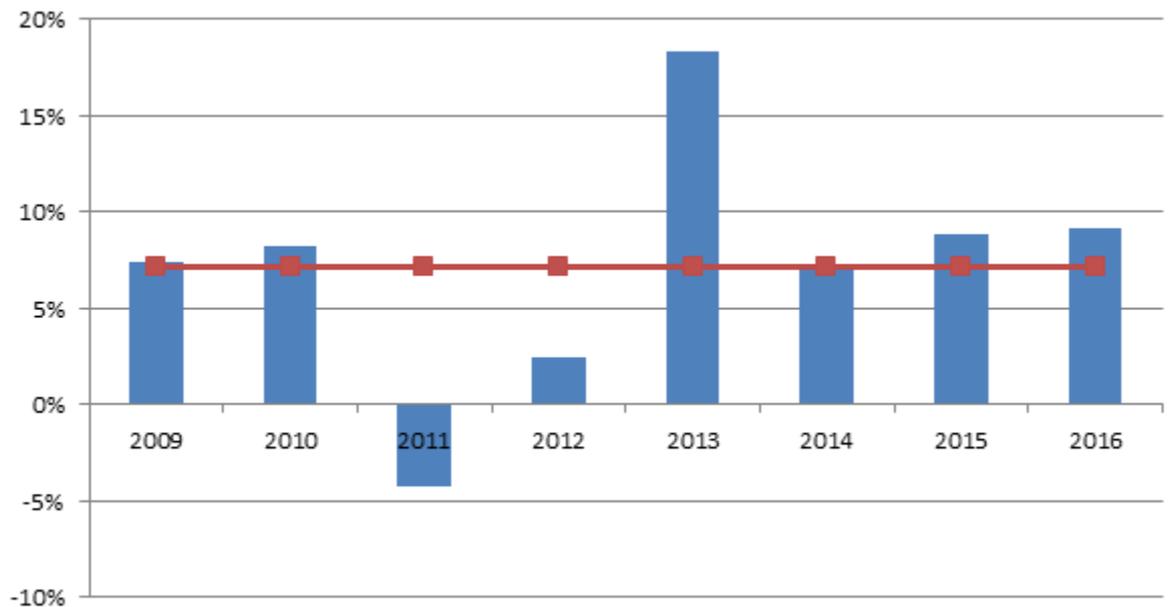
Проблемы использования

=СрЗнач()

Выборочное среднее (компания А)

$$\overline{X}_A = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{7,4\% + 8,2\% + \dots + 9,1\%}{8} = 7,15\%$$

Год	\overline{X}_A	Компания А
2009	$X_1 =$	7,4 %
2010	$X_2 =$	8,2 %
2011	$X_3 =$	- 4,3 %
2012	$X_4 =$	2,4 %
2013	$X_5 =$	18,3 %
2014	$X_6 =$	7,3 %
2015	$X_7 =$	8,8 %
2016	$X_8 =$	9,1 %



Выборочное среднее (компания Б)

$$\overline{X}_B = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{6,2\% + 9,7\% + \dots + 9,7\%}{8} = 7,26\%$$

Год	\overline{X}_B	Компания Б
2009	$X_1 =$	6,2 %
2010	$X_2 =$	9,7 %
2011	$X_3 =$	- 5,8 %
2012	$X_4 =$	3,5 %
2013	$X_5 =$	20,6 %
2014	$X_6 =$	4,5 %
2015	$X_7 =$	9,7 %
2016	$X_8 =$	9,7 %



Выборочное среднее (сравнение компаний А и Б)

Компания А: 7,15 %

Компания Б: 7,26 %

Вывод: средняя рентабельность компании Б выше!

Влияние выбросов на результаты оценок

$$\bar{X}_A = \frac{X_1 + \dots + X_n}{n} = \frac{7,4\% + 8,2\% + \dots + 9,1\%}{7} = 5,56\%$$

$$\bar{X}_B = \frac{X_1 + \dots + X_n}{n} = \frac{6,2\% + 9,7\% + \dots + 9,7\%}{7} = 5,36\%$$

Какую
компанию
выбрать?

Год	Компания А	Компания Б
2009	7,4 %	6,2 %
2010	8,2 %	9,7 %
2011	- 4,3 %	- 5,8 %
2012	2,4 %	3,5 %
2013	18,3 %	20,6 %
2014	7,3 %	4,5 %
2015	8,8 %	9,7 %
2016	9,1 %	9,7 %

Медиана и мода

Медиана (M) – число, разделяющее выборку пополам: 50% элементов меньше медианы, а 50% больше

Если число элементов в выборке нечетно

$$M = X_{\frac{n+1}{2}}$$

Если число элементов в выборке четно

$$M = \left(X_{\frac{n}{2}} + X_{\frac{n}{2}+1} \right) / 2$$

Мода (K) – число, чаще других встречающееся в выборке

Проблемы использования

1. Мода сильно зависит от конкретной выборки
2. Мода не всегда существует



Выборку
необходимо
упорядочить!

Медиана и мода (компании А и Б)

1. Медиана рентабельности компании **А** составляет **7,80%**
2. Мода по компании **А** отсутствует

3. Результаты по компании Б:

Год	\bar{X}_B	Компания Б
2009	X ₁ =	6,2 %
2010	X ₂ =	9,7 %
2011	X ₃ =	-5,8 %
2012	X ₄ =	3,5 %
2013	X ₅ =	20,6 %
2014	X ₆ =	4,5 %
2015	X ₇ =	9,7 %
2016	X ₈ =	9,7 %

Год	Рентабельность (%)
2011	-5,8
2012	3,5
2014	4,5
2009	6,2
2010	9,7
2015	9,7
2016	9,7
2013	20,6

Массив упорядочен по возрастанию

$$M = \frac{6,2 + 9,7}{2} = 7,95$$

$$K = 9,70$$

=Медиана()

=Мода()

Сравнение значений показателей (1)

Наименование компании	Выборочное среднее	Выборочное среднее*	Медиана	Мода
Компания А	7,15	<u>5,56</u>	7,80	-
Компания Б	<u>7,26</u>	5,36	<u>7,95</u>	<u>9,70</u>

Выводы?

Медиана или выборочное среднее – что выбрать?

Например: производится анализ средней урожайности (X_y), которая в прошлом изменялась так, как показано ниже (тыс. тонн). Т.е. значение урожайности - достаточно стабильный показатель, но каждые 3-4 года бывают «существенные» неурожаи

2001	2002	2003	2004	2005	2006	2007	2008	2009
10,5	10,1	10,8	5,4	11,0	9,5	4,5	10,3	10,0

Выборочное среднее = 9 тыс. тонн; медиана = 10 тыс. тонн.

А для каких целей производится расчет ???

Задача: прогноз по валовому сбору на ближайшие 10 лет



Среднее арифметическое

Задача: понять, какова урожайность сорта при благоприятных условиях



Медиана

Исследование числовых данных (3)

Второе важное свойство числовых данных – их **вариация** (т.е. разброс).

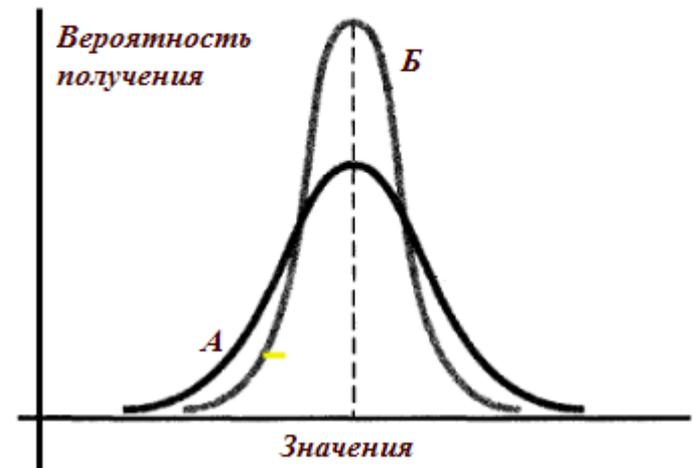
Рассмотрим четыре числовые характеристики, используемые для оценки разброса значений:

- размах;
- дисперсия;
- стандартное отклонение;
- вариация

Одинаковый разброс, но разные средние значения



Одинаковые средние значения, но разный разброс



Размах (1)

Размах (Z) – разность между наибольшим и наименьшим элементами выборки:

$$Z = X_{Max} - X_{Min}$$

Проблемы использования



Размах (2)

Оцените размах значений рентабельности компаний А и Б

Год	Компания А	Компания Б
2009	7,4 %	6,2 %
2010	8,2 %	9,7 %
2011	- 4,3 %	- 5,8 %
2012	2,4 %	3,5 %
2013	18,3 %	20,6 %
2014	7,3 %	4,5 %
2015	8,8 %	9,7 %
2016	9,1 %	9,7 %

$$Z_A = 18,3\% - (-4,3\%) = 22,6\%$$

$$Z_B = 20,6\% - (-5,8\%) = 26,4\%$$

Вывод: разница между максимальным и минимальными значениями рентабельности компании Б больше, следовательно рентабельность второй компании имеет больший разброс

Дисперсия и стандартное отклонение (1)

Дисперсия (D) и стандартное (среднеквадратическое) отклонение (σ) – меры разброса случайной величины, то есть её отклонения от среднего значения, учитывающие как именно распределены данные:

$$D = \frac{[(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2]}{(n-1)} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}$$

$$\sigma = \sqrt{D}$$

Этапы вычисления показателей дисперсии и СКО:

1. Вычислить разность между каждым элементов выборки (X_i) и выборочным средним (\bar{X})
2. Возвести каждую разность в квадрат
3. Сложить все разности, возведенные в квадрат
4. Поделить результат на $n - 1$
5. Извлечь квадратный корень из выборочной дисперсии

Дисперсия и стандартное отклонение (2)

Компания А

Рентабельность (%)	\bar{X} Среднее значение	$(X_i - \bar{X})$ Разность между средним и элементом выборки	$(X_i - \bar{X})^2$ Возведение каждой разности в квадрат
7,4	7,15	0,25	0,0625
8,2	7,15	1,05	1,1025
-4,3	7,15	-11,45	131,1025
2,4	7,15	-4,75	22,5625
18,3	7,15	11,15	124,3225
7,3	7,15	0,15	0,0225
8,8	7,15	1,65	2,7225
9,1	7,15	1,95	3,8025
Сумма:			285,70

СКО (σ) = 6,39%

Компания Б

СКО (σ) = 7,46%

=СТАНДОТКЛОН.В()

Интерпретация дисперсии и стандартного отклонения (1)

Вычисленное СКО(А) = 6,39%. Рентабельность большинства периодов компании А лежит в пределах среднего значения ± 1 СКО, т.е. [0,76%; 13,54%]. Фактически в этом интервале лежат значения рентабельности 6 из 8 периодов.

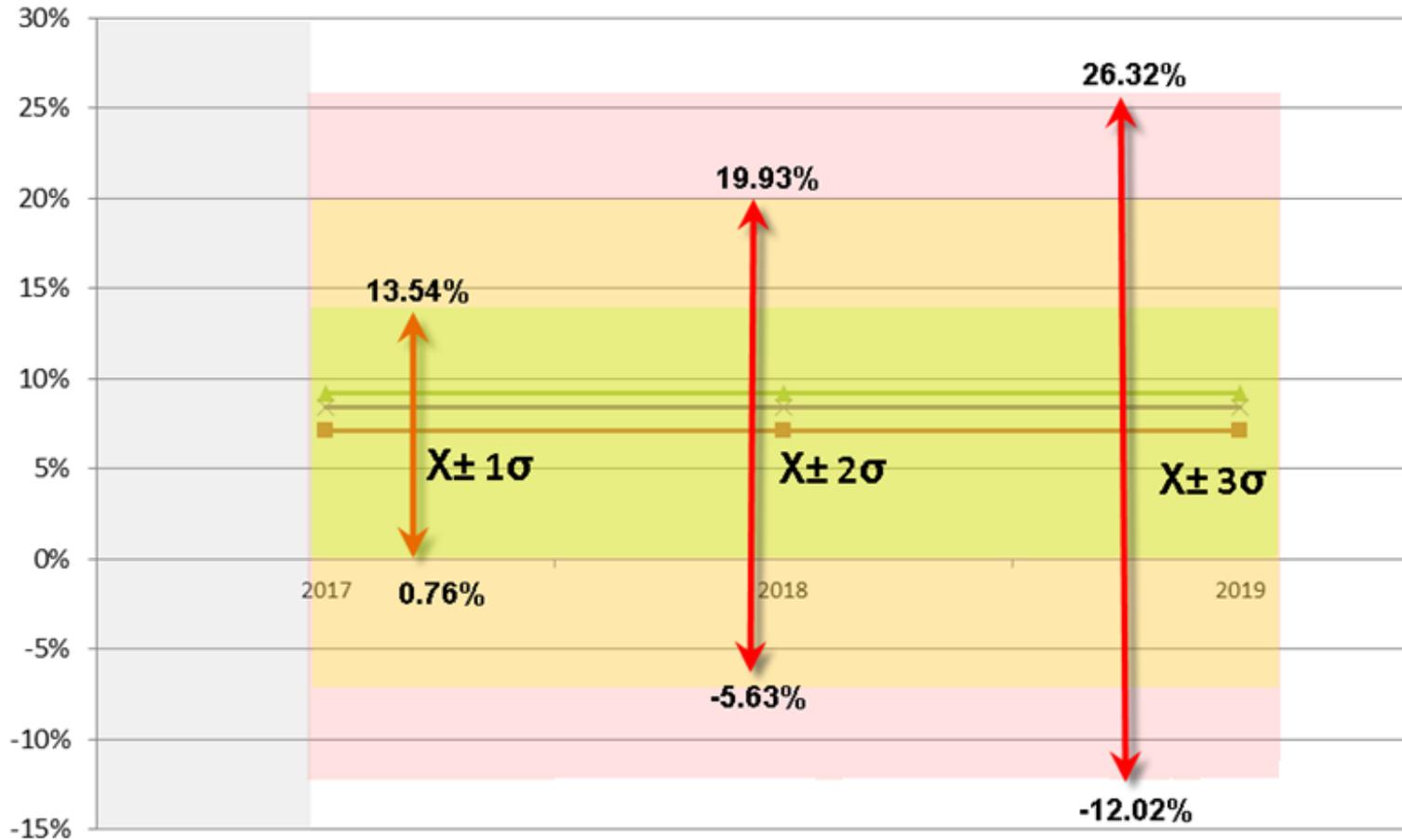
Т.е. зная выборочное среднее и СКО выборки можно определить интервал, которому принадлежит основная масса данных!

Год	Рентабельность (%)
2009	7,4
2010	8,2
2011	-4,3
2012	2,4
2013	18,3
2014	7,3
2015	8,8
2016	9,1

Дисперсия представляет собой квадрат ед. изм.!

Вычисление диапазона будущих значений (2)

Компания А



Оцените самостоятельно диапазон значений для компании Б!!

Правило Бьенаме-Чебышева

Более ста лет назад два математика - Бьенаме и Чебышев - обнаружили, что независимо от формы распределения данных, процент наблюдений, лежащий на расстоянии, не превышающим k стандартных отклонений от среднего значения (математического ожидания – для всей совокупности данных), не меньше, чем:

$$\left(1 - 1/k^2\right) * 100\%$$

Процент наблюдений, попадающий в интервал		
<i>Интервал</i>	<i>Правило Б-Ч</i>	<i>Эмпирическое правило*</i>
$\pm\sigma$	как минимум 0%	около 68%
$\pm 2\sigma$	как минимум 75%	около 95%
$\pm 3\sigma$	как минимум 88,89%	около 99,7%

Коэффициент вариации

Вариация (CV) – риск, приходящийся на единицу доходности:

$$CV = \frac{\sigma}{X}$$

Компания А

$$\overline{X}_A = 7,15\%$$

$$\sigma_A = 6,39\%$$

$$CV = 6,39\% / 7,15\% = 0,89$$

Компания Б

$$\overline{X}_B = 7,26\%$$

$$\sigma_B = 7,46\%$$

$$CV = 7,46\% / 7,26\% = 1,03$$

Вывод: Компания Б в 1,15 раз более рискованна чем компания А

Вероятность....

Возможность того, что какое-то событие или внешнее условие наступит, известно под термином **вероятность**.

Вероятность выражается в виде десятичной дроби, при этом ее значение находится между 0 и 1, или в %:

0 – событие не произойдет;

1 – событие обязательно случится.

Вероятность 0,4 означает, что событие случится 4 раза из 10.

Например: вероятность успешно сдать курс составляет 0,7, следовательно вероятность не сдать курс составляет 0,3.

Распределение вероятностей

Исход	Вероятность
Сдача курса	0,7
Провал курса	0,3
Итого	<u>1,0</u>

Объективные и субъективные вероятности

Объективные вероятности – события, оцененные математически, или полученные на основании исторически достоверных данных.

Субъективные вероятности – события, вероятность которых оценена менеджментом на основе собственных суждений.

- + Получены на основе прошлого опыта
- Могут возникать субъективные ошибки

Метод анализа иерархий
(АНР)

Расчет средних и отклонений на основе ожидаемых значений

$$\bar{k} = \sum_{i=1}^m P_i * k_i \quad \sigma = \sqrt{\sum_{i=1}^m P_i (k_i - \bar{k})^2}$$

Компания А

Год	Доверие	-	Значения	Среднее	Отклонение	Отклонение^2	P*Отклонение^2	Сумма (дисперсия)	СКО
2009	0,05	X ₁ =	7,40%	7,99%	-0,00590	0,00003	0,00000	0,00151	0,03883
2010	0,05	X ₂ =	8,20%		0,00210	0,00000	0,00000		
2011	0,05	X ₃ =	-4,30%		-0,12290	0,01510	0,00076		
2012	0,05	X ₄ =	2,40%		-0,05590	0,00312	0,00016		
2013	0,05	X ₅ =	18,30%		0,10310	0,01063	0,00053		
2014	0,2	X ₆ =	7,30%		-0,00690	0,00005	0,00001		
2015	0,25	X ₇ =	8,80%		0,00810	0,00007	0,00002		
2016	0,3	X ₈ =	9,10%		0,01110	0,00012	0,00004		

Оцените самостоятельно СКО для компании Б (уровень доверия к данным по годам тот же, что и для компании А)!!

Выводы?

Кейс

Предположим, что Бриджет Нельсон, президент инвестиционной компании, получает предложение об участии в ряде проектов. Экспертами компании были спрогнозированы вероятные тенденции развития рынка на ближайшие 5 лет, совпадающие с периодами реализации проектов. Выполненные расчеты позволили оценить итоговую доходность этих проектов при различных прогнозах развития рынка (табл. 1).

К ближайшему совету директоров Вам необходимо:

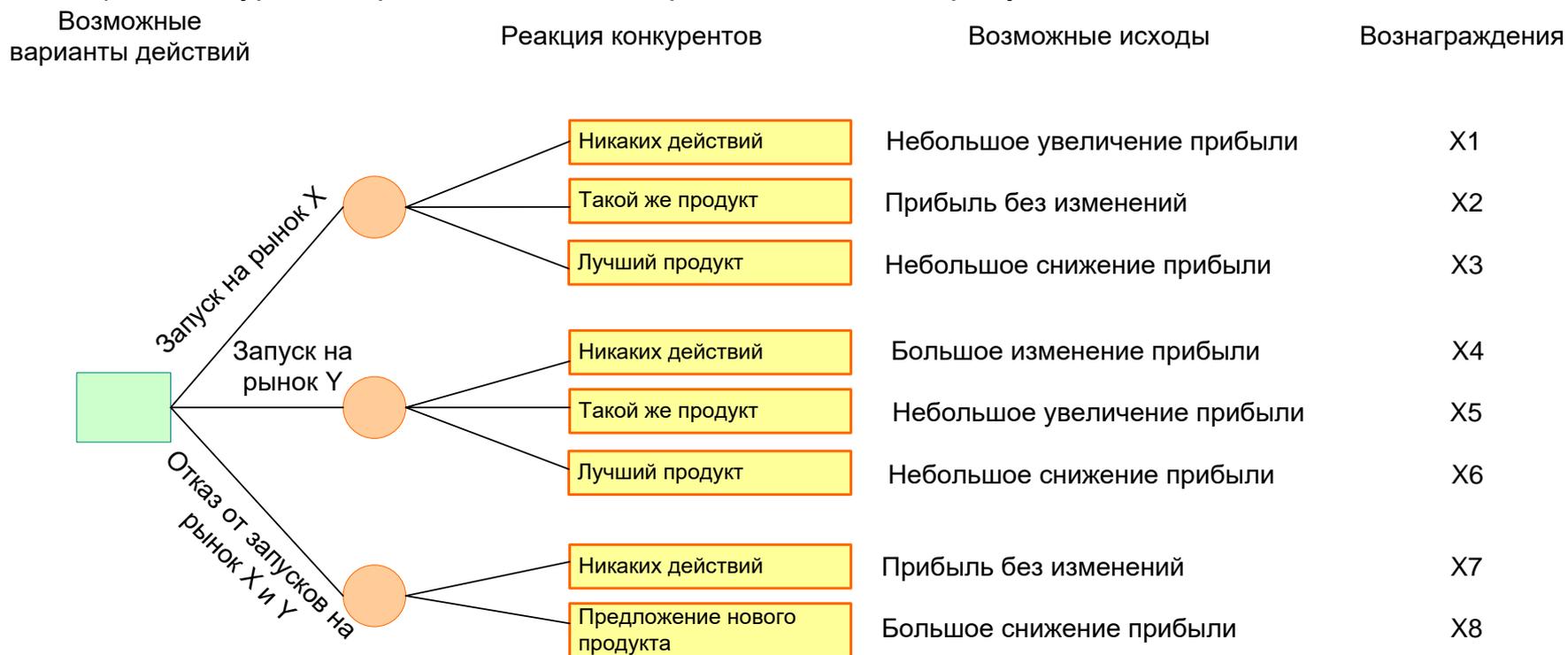
1. Выбрать проект в котором наиболее предпочтительно участие компании
2. Оценить среднеквадратическое отклонение доходности каждого из проектов и провести интерпретацию полученных результатов.
3. Сделать общие выводы о рискованности предлагаемых инвестиций

Тенденция рынка	Вероятность	Проект «А»	Проект «Б»	Проект «В»
Рыночный бум	60%	30%	20%	14%
Средний спрос на продукцию	15%	10%	16%	10%
Отсутствие спроса	25%	4%	9%	8%

Дерево решений

Pretorian Company рассматривает свою маркетинговую политику на следующий период. Компания разработала два новых продукта – X и Y, но ее ресурсов достаточно для выпуска только одного из них. С точки зрения конкурентов компания может столкнуться с одним из следующих действий:

- 1) конкуренты ничего не предпринимают;
- 2) конкуренты предлагают такой же продукт;
- 3) конкуренты предлагают более привлекательный продукт.



Задача

Компания прогнозирует целесообразность разработки нового продукта и выхода с ним на рынок. Затраты на разработку 180 000 \$, и существует вероятность 0,75, что разработки будут успешными и 0,25, что окончатся неудачей. Если разработки будут успешными, продукт будет выведен на рынок, при этом существуют следующие оценки:

Результат	Вероятность	Прибыль (\$)*
Высокая успешность	0,4	540 000
Средняя успешность	0,3	100 000
Провал	0,3	- 400 000

1. Изобразите дерево решений
2. Оцените целесообразность разработки продукта
3. Оцените итоговые вероятности убытков и прибыли компании

Задача

Банк решает вопрос, проверять ли конкурентоспособность клиента перед тем, как выдавать ему кредит. Аудиторская фирма берет 80\$ за каждую проверку. В результате перед банком встают две проблемы:

- 1) доверять ли в будущем проверку аудиторам;
- 2) выдавать ли кредит при отрицательных результатах экспертизы.

Решая первую проблему, банк проверяет правильность сведений аудиторской фирмы. Для этого выбираются 1000 человек, которым выдавались ссуды (табл. ниже).

Банк выдает фиксированные кредиты сроком на один год по 15 000\$, под 15% годовых. Альтернатива - вложить деньги «в дело» со 100%-ным возвратом суммы, но под 9% годовых

Рекомендации аудиторской фирмы	Фактический результат		
	Клиент вернул ссуду	Клиент не вернул ссуду	Всего
Давать кредит	735	15	750
Не давать кредит	225	25	250
ВСЕГО	960	40	1000

Задачи: приобретение информации

Компания ВС выбирает между двумя типами оборудования: оборудование А и В. Пусть существует два возможных варианта спроса низкий и высокий, причем вероятность каждого из них 0,5.

Оценки прибыли для каждого спроса следующие:

	Низкий спрос (\$)	Высокий спрос (\$)	Ожидаемое значение (\$)
Оборудование А	100 000	160 000	130 000
Оборудование В	10 000	200 000	105 000

Компания может пригласить консультантов, которые смогут точно вычислить итоговый спрос. Сколько денег можно заплатить консультантам за информацию?

Критерий максимин, максимакс

Критерий МаксиМин. Прогнозирует, что исходом процесса будет самый плохой вариант, и поэтому следует принять решение, которое обеспечивает лучший результат в худшем варианте. В прошлом примере необходимо выбрать оборудование А, так как при низком спросе оно дает лучший результат.

Критерий МаксиМакс. Прогнозирует, что исходом процесса будет самый хороший вариант, и поэтому следует принять решение, которое обеспечивает лучший результат в лучшем варианте. В прошлом примере необходимо выбрать оборудование Б, так как при высоком спросе оно дает лучший результат.

Критерий сожаления. Прогнозирует, что необходимо минимизировать потери в следствии неправильно выбранного варианта. Следовательно необходимо выбрать оборудование А, т.к. итоговые потери меньше.

	Низкий спрос (\$)	Высокий спрос (\$)
Оборудование А	0	40 000
Оборудование В	90 000	0

Ковариация

Ковариация (Cov) – показатель, учитывающий тенденцию к росту или снижению одного показателя по мере того как изменяется другой:

$$Cov(AB) = \sum_{i=1}^n (k_{Ai} - \bar{k}_A)(k_{Bi} - \bar{k}_B)P_i$$

Свойства

1. Если коэффициенты K_{Ai} , K_{Bi} меняются в одном направлении относительно своих средних значений, то оба выражения в скобках будут иметь один знак и наоборот. Т.е. если K_{Ai} , K_{Bi} изменяются сонаправлено, то Cov – положительна и наоборот, а если изменяются случайным образом, то $Cov \rightarrow 0$
2. Если K_{Ai} , K_{Bi} претерпевают большие колебания, то значения в скобках будут большими и Cov будет увеличиваться, однако Cov может оказаться и малой, если отрицательные и положительные члены компенсируют друг друга
3. Если величина имеет нулевое σ , то все ее отклонения равны нулю и Cov будет также нулевой

Ковариация изменяется от $(-\infty; +\infty)$

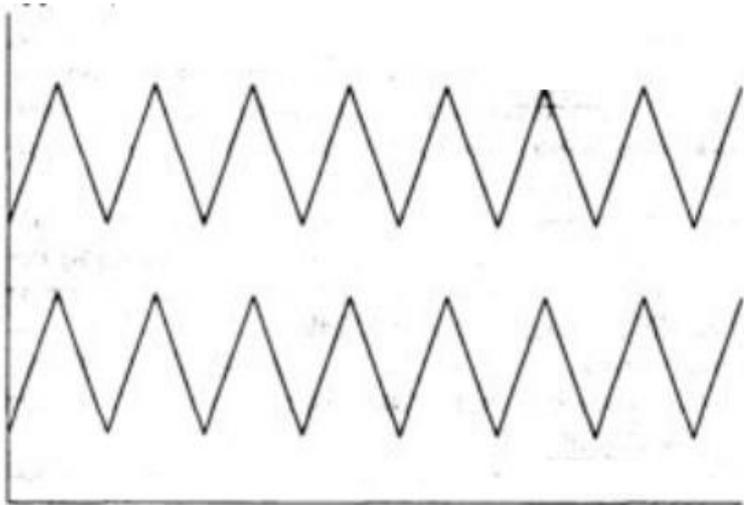
Коэффициент корреляции

Корреляция (r_{AB}) – показатель, который стандартизует ковариантность, что облегчает сравнения переменных:

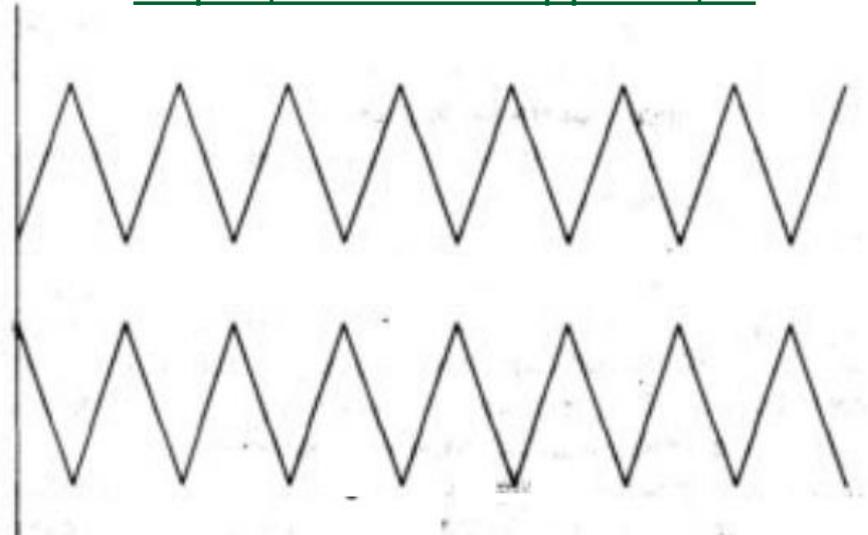
$$r_{AB} = Cov(AB) / \sigma_A \sigma_B$$

Корреляция изменяется от [-1; + 1]

Положительная корреляция



Отрицательная корреляция



Задача

Прогнозируется, что акции компаний А и Б будут иметь следующие доходности:

Вероятность	Доходность акций А, K_a %	Доходность акций В, K_b %
0,2	-10%	-4%
0,3	18%	24%
0,5	14%	15%

Оцените ковариацию и корреляцию доходности акций

Корреляция: практический пример

OIL PRICES BY QUARTER \$/bbl REAL 2010 PRICES

	Brent	WTI	JCC
Q1 2005	54,5	50,2	41,0
Q2 2005	58,7	53,1	49,7
Q3 2005	69,5	63,2	56,0
Q4 2005	63,7	60,0	57,4
Q1 2006	68,6	63,3	59,5
Q2 2006	76,7	70,4	66,9
Q3 2006	76,0	70,4	70,6
Q4 2006	64,7	60,0	61,0
Q1 2007	62,0	58,1	57,6
Q2 2007	73,3	65,0	64,8
Q3 2007	79,3	75,5	71,0
Q4 2007	93,2	90,8	82,8
Q1 2008	100,8	98,0	93,1
Q2 2008	125,6	124,0	110,2
Q3 2008	118,2	118,0	129,1
Q4 2008	56,2	58,4	77,0
Q1 2009	45,0	43,0	44,3
Q2 2009	59,4	59,5	53,0
Q3 2009	68,8	68,2	70,3
Q4 2009	75,0	76,1	74,8
Q1 2010	76,8	79,2	78,2
Q2 2010	78,5	78,0	81,7
Q3 2010	76,7	75,9	75,5

REGIONAL COAL PRICES

	NW Europe \$/mmbtu	Japan \$/mmbtu
	1,41	2,26
	2,28	3,13
	2,33	3,18
	2,39	3,24
	2,30	3,15
	2,13	2,98
	2,15	3,00
	2,13	2,98
	2,22	3,07
	2,19	3,04
	2,31	3,16
	2,45	3,30
	2,96	3,81
	3,50	4,35

	Brent	WTI	JCC	NW Europe \$/mmbtu	Japan \$/mmbtu
Brent	-	0,98	0,98	0,58	0,58
WTI	-	-	0,96	0,54	0,54
JCC	-	-	-	0,63	0,64
NW Europe \$/mmbtu	-	-	-	-	0,94
Japan \$/mmbtu	-	-	-	-	-

Выводы:

1. Для будущих прогнозов только значения на одном из рынков нефти или угля, можно предсказать поведение цен на другом рынке
2. Наблюдается слабая зависимость между ценами на уголь и нефть. Таким образом, можно говорить о слабой взаимозаменяемости этих энергоресурсов

СПАСИБО ЗА ВНИМАНИЕ!

Литвин Юрий Васильевич

Email: litvinj@simplecs.ru

Тел.: 8 (929) 906-90-33