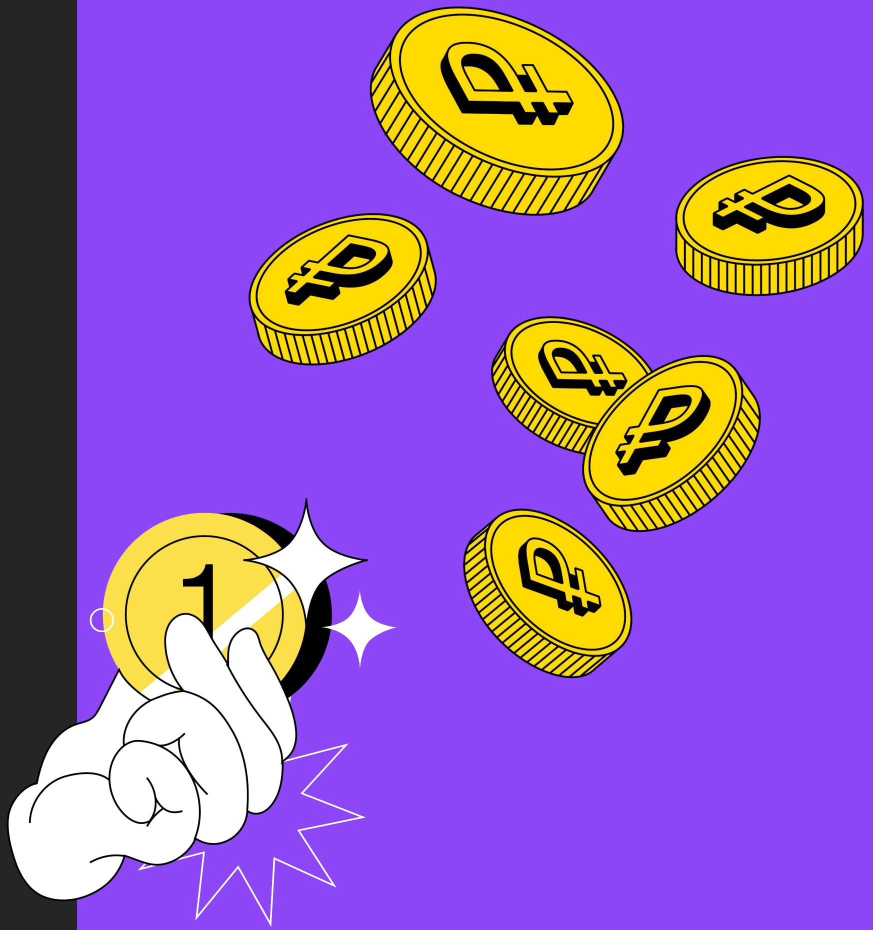
GeekBrains

Применение статистики в экономике и финансах

Урок 4



Знакомство и содержание урока

Дмитрий Бородин

Инвестиционный директор венчурного фонда

- 💥 **12 лет опыта** в корпоративных финансах, инвестиционной оценке и финансовом моделировании.
- 💥 **14 лет опыта** преподавания математических и финансово-экономических дисциплин.
- 💥 Кандидат экономических наук, доцент **МГТУ им. Н.Э. Баумана**, приглашенный преподаватель РЭУ им. Плеханова.
- 💥 Сертификаты финансового аналитика **CFA** и финансового риск-менеджера **FRM.**

План курса

1. Временная стоимость денег
2. Проведение процентных расчетов
3. Оценка эффективности инвестиций
4. Применение статистики в экономике и финансах
5. Введение в эконометрику

Что будет на уроке сегодня

- 📌 Познакомимся с основными терминами теории вероятностей и математической статистики.
- 📌 Рассмотрим применение статистических концепций в экономике и финансах.
- 📌 Познакомимся с ролью риска и неопределённости в современных финансах.
- 📌 Изучим расчёт доходности и риска финансовых активов.
- 📌 Познакомимся с акциями и разберём их роль на финансовых рынках.
- 📌 Рассмотрим основные теории и формулы портфельного анализа.
- 📌 Научимся применять теорию вероятностей и математическую статистику для практического анализа инвестиционных портфелей.

Введение в теорию вероятностей

Что такое теория вероятностей?

Теория вероятностей — раздел математики, изучающий случайные события, случайные величины и их свойства.

Теория вероятностей — это база для математической статистики, которая, в свою очередь, лежит **в основе** современной финансовой теории риска и доходности.

Основные термины теории вероятностей

Случайное событие – событие, у которого нет заранее известного исхода. Набор вероятных исходов случайного события всегда больше одного.

Случайная величина – величина, чьё значение меняется в зависимости от случая (т.е. неопределённости). Случайная величина может принимать одно из нескольких значений, каждое с некоторой ассоциированной вероятностью.

Случайные величины могут быть:

- **дискретными** – принимают одно из определённого набора значений,
- **непрерывными** – принимают любое значение внутри некоторого интервала.

Вероятность – мера возможности наступления некоторого события. Определяется как число в интервале от 0 до 1 (где 0 – абсолютная невозможность, а 1 – достоверность). Чем выше вероятность события, тем более мы уверены в том, что событие произойдёт. Вероятность события X обычно обозначается как P(X).

Знание вероятности позволяет найти **ожидаемое значение**:

$$E(X) = X1*P(X1) + X2*P(X2) + \dots$$

Основные формулы теории вероятностей

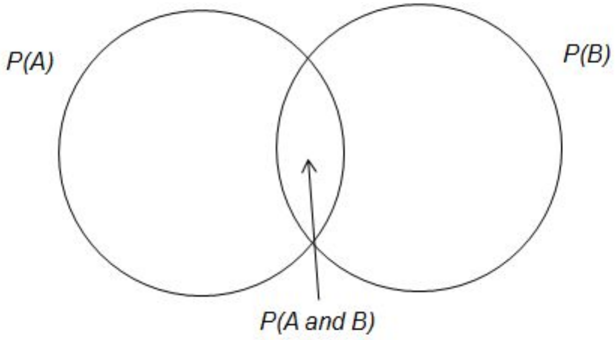
Вероятность события: ***P(A) = M / N***

Вероятность противоположного события: ***P(не A) = 1 – P(A)***

Независимые события: $P(A \text{ and } B) = P(A \cap B) = P(A)P(B)$

Несовместные события: $P(A \text{ or } B) = P(A \cup B) = P(A) + P(B)$

Сложение вероятностей: $P\,(A \text{ or } B) = P\,(A) + P\,(B) - P\,(A \text{ and } B)$



Условные вероятности и формула Байеса

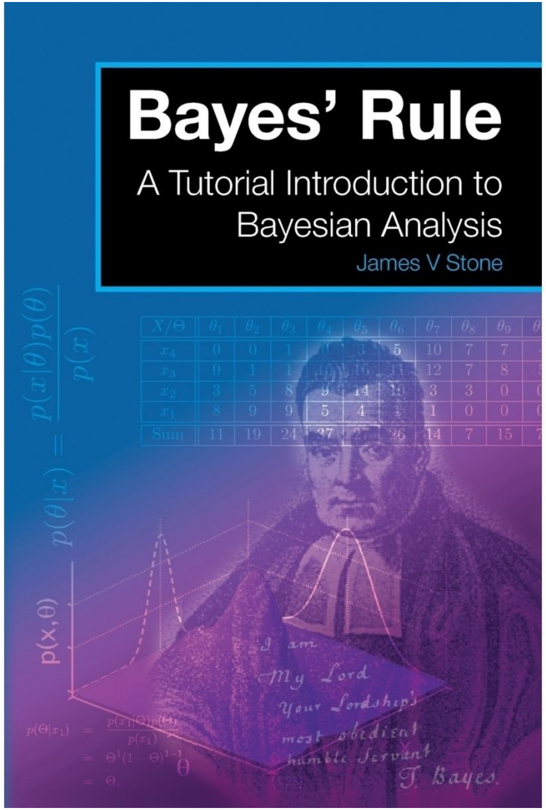
Условная вероятность – вероятность некоторого события А при условии наступления некоторого события *B:*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Теорема Байеса описывает вероятность события А на основе информации о некоторых условиях В, которые связаны с наступлением события :

$$P(A|B) = \frac{P(A)\,P(B|A)}{P(B)}$$

Рассмотрим использование формул теории вероятности на практических примерах.



Примеры использования формул

Задача 1. В городе 85% автомобилей зелёные и 15% – синие. Свидетель ДТП ночью идентифицировал автомобиль как синий. Однако известно, что в сумерках человек правильно идентифицирует цвет в 80% случаев. Какова вероятность, что автомобиль-нарушитель был действительно синий?

Решение

Событие **A** – «автомобиль синий», P(A) = 15%;

Событие **B|A** – автомобиль идентифицирован как синий, в то время как он на самом деле синий, P(B|A) = 80%;

Событие **B** – «автомобиль идентифицирован как синий», P(B) = 80% * 15% + 20% * 85% = 29%.

Ответ: P(A|B) = 15% * 80% / 29% = 41%.

Примеры использования формул

Задача 2. Если известны условные вероятности проведения акции в разных состояниях экономики, и известно, что акция выросла, какова вероятность, что состояние экономики позитивное?

Решение. Общая вероятность, что акция выросла = 30%*60% + 50%*30% + 20%*10% = 35%, а нас интересует только ситуация, когда акция выросла при позитивной экономике = 30%*60% = 18%.

Ответ: 18%/35% = 51%.

Состояние экономики страны	Вероятность этого состояния	Поведение акции	Условная вероятность
Позитивное	30%	Рост	60%
		Нейтрально	30%
		Падение	10%
Нейтральное	50%	Рост	30%
		Нейтрально	40%
		Падение	30%
Негативное	20%	Рост	10%
		Нейтрально	60%
		Падение	30%

Введение в математическую статистику

Что такое математическая статистика?

Математическая статистика – раздел математики, изучающий методы описания и анализа данных наблюдений и экспериментов с целью прогнозирования случайных событий.

Математическая статистика изучает исторические значения случайной величины и позволяет на основе их анализа делать прогнозы о будущих значениях => современные финансы полностью основаны на математической статистике!

Основные термины математической статистики

Все изучаемые случайные величины делятся на дискретные и непрерывные.

- Для каждого значения **дискретной случайной величины** можно однозначно сопоставить вероятность появления такого значения, поскольку их число конечно и известно.
- Для **непрерывной случайной величины** нет возможности сопоставить вероятность каждому значению (поскольку число возможных значений бесконечно), можно лишь говорить о вероятности попадания в определённый интервал значений.

Генеральная совокупность – набор всех возможных значений, которые случайная величина принимала за всё время наблюдения (например, доходность индекса S&P500 с момента его появления до сегодняшнего дня).

Выборка – некоторое подмножество исторических значений случайной величины (например, доходность индекса S&P500 за период 2015-2020 гг).

Основные формулы математической статистики

Функция вероятности p(x) — зависимость вероятности появления определённого значения «x» случайной величины X от этого значения:

$$p(x) = P(X=x)$$

Функция вероятности всегда ограничена 0 и 1:

$$0 \le p(x) \le 1$$

Поскольку случайная величина гарантированно примет одно из набора её значений, то сумма вероятностей для всех значений всегда равна 1:

$$\Sigma p(x) = 1$$

Среднее значение случайной величины по историческим данным — среднее арифметическое всех значений, входящих в рассматриваемый период. Часто среднее значение называют **математическим ожиданием** или **матожиданием**. Матожидание может быть рассчитано как по генеральной совокупности, так и по выборке:

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N} \qquad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Основные формулы математической статистики

Медиана — средняя величина в упорядоченном ряду исторических значений. Преимущество медианы — на неё не влияют экстремальные исторические значения величины («выбросы»).

Мода — то значение случайной величины, которое чаще всего встречается в историческом наборе данных.

Геометрическое среднее рассчитывается по формуле: $G = \sqrt[n]{X_1 \times X_2 \times \ldots \times X_n} = (X_1 \times X_2 \times \ldots \times X_n)^{1/n}$

Гармоническое среднее рассчитывается по формуле: $H = \dfrac{N}{\sum_{i=1}^{N} \frac{1}{X_i}}$

Дисперсия — мера разброса значений случайной величины относительно её математического ожидания. Дисперсия генеральной совокупности и дисперсия выборки рассчитываются по формулам:

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N} \qquad s^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}$$

Среднеквадратичное отклонение (СКО) — квадратный корень из дисперсии.

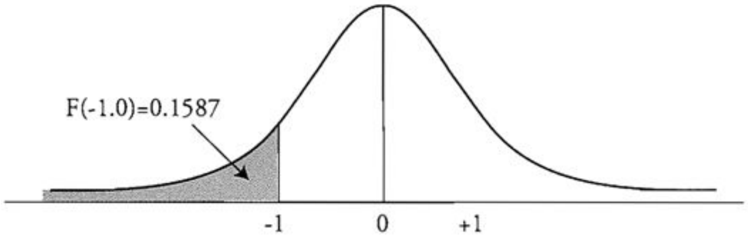
Законы распределения случайных величин

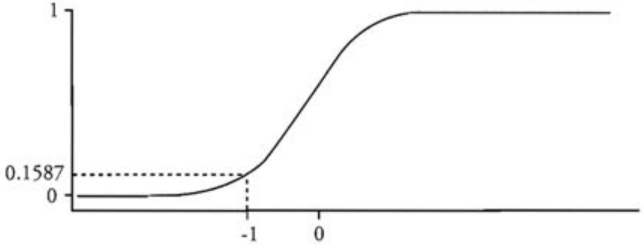
Для непрерывных случайных величин функция вероятности p(x) не имеет смысла, поскольку возможных значений бесконечное множество, и вероятность одного конкретного значения всегда равна 0!

Вместо этого используется **функция плотности вероятности f(x)** — она позволяет определить вероятность попадания случайной величины в определённый интервал.

Альтернативный способ описания распределения случайной величины — с помощью (**кумулятивной) функции распределения F(x)**, которая показывает вероятность того, что случайная величина X меньше или равна конкретному значению x:

$$F(x) = P(X \leq x)$$

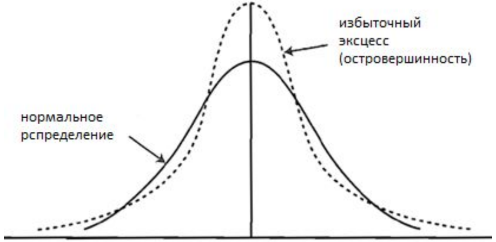




Описание распределения случайной величины

1. Матожидание (среднее значение).
2. Среднеквадратичное отклонение или дисперсия.
3. Симметричность распределения относительно среднего значения.

1. Эксцесс или островершинность распределения обычно измеряется как избыточный эксцесс относительно нормального распределения.

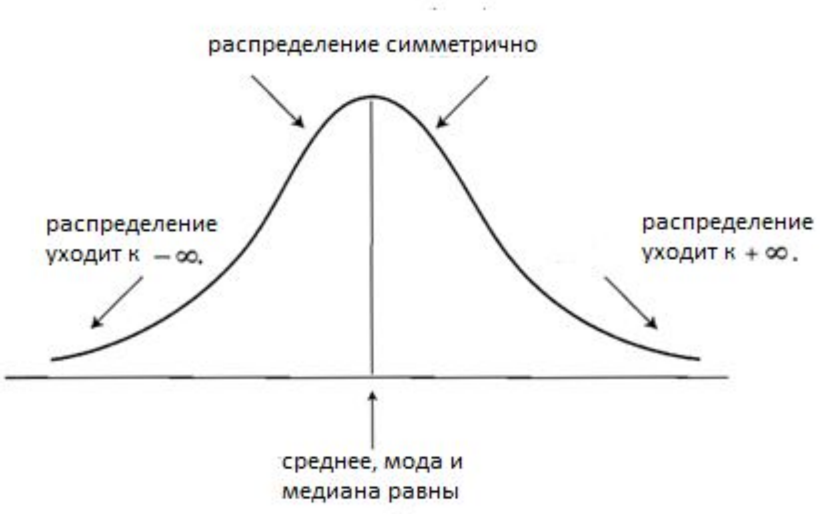


Нормальное распределение и центральная предельная теорема

Нормальное распределение (распределение Гаусса) — это наиболее используемое распределение случайной величины в финансово-экономических расчётах.

Центральная предельная теорема (ЦПТ): для произвольной выборки из **n** значений случайной величины с матожиданием **μ** и дисперсией **σ^2**, распределение среднего значение по этой выборке стремится к нормальному распределению с матожиданием **μ** и дисперсией **σ^2 / n** по мере того, как растёт размер выборки **n**.

Другое название ЦПТ — **закон больших чисел**. Широко используется, поскольку нормальное распределение легко применять для практических расчётов: вне зависимости от реального распределения случайной величины для больших размеров выборки с ней можно работать как с нормальным распределением (на практике для n > 30).



$$z = \frac{\text{значение} - \text{матожидание}}{\text{СКО}} = \frac{x-\mu}{\sigma} \sim N(0,1)$$

Риск и доходность

Риск и доходность финансовых активов

Доходность активов на финансовых рынках — процент от стоимости актива, который инвестор заработал в результате инвестиции. Для торгуемых на бирже финансовых активов принято рассчитывать доходность за период по формуле:

$$R(i) = \ln\left(P(i)/P(i-1)\right)$$

Обратите внимание, что это не арифметическая доходность, а так называемая геометрическая доходность!

Формула применяется только для малых периодов (минута, час, день). Если нужно определить доходность за длительный период (месяц, год), мы должны пересчитать из среднедневной доходности: $R_{год} = R_{дн} * 252$ (число торговых дней).

Моделирование цен акций

Акция — ценная бумага, подтверждающая владение частью компании (бизнеса), право на получение дивидендов и доли имущества в компании при ликвидации. Акции — один из основных инструментов фондового рынка (наряду с облигациями). Составляют существенную часть инвестиционных портфелей инвесторов по всему миру.

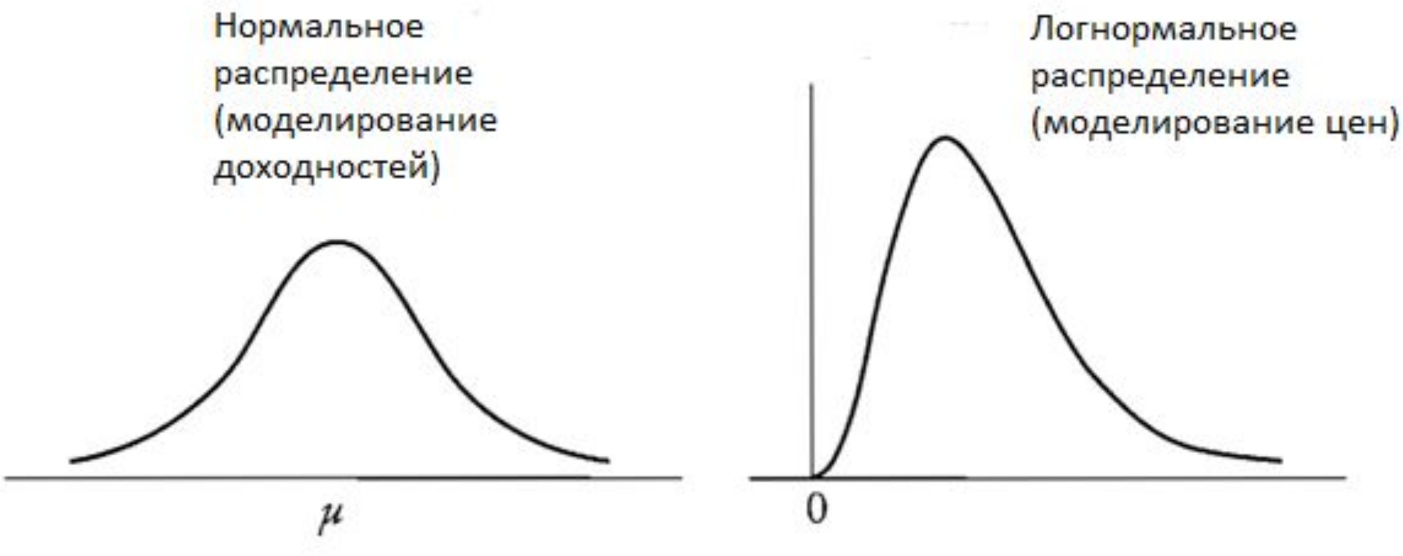
В отличие от облигаций, где денежный поток известен наперёд и гарантирован, денежный поток по акциям неизвестен и представляет собой **дивиденды** (регулярно выплачиваемую часть прибыли компании) и **изменение цены акции** (доход инвестора равен цене продажи минус цена покупки).

Поэтому для анализа инвестиций в акции приходится прибегать к статистике — анализировать историческую динамику цены акции, изучать её статистические характеристики и делать предположения на будущее.

Моделирование цен акций

Нормальное распределение не может быть использовано для моделирования цен, потому что цены ограничены 0.

Поэтому нормальным распределением моделируются доходности активов, а их цены моделируются так называемым **логнормальным распределением**:



Риск и неопределённость

Знания средней ожидаемой доходности акции не хватает для принятия решения об инвестировании в неё, поскольку всегда есть **неопределённость** того, что произойдет в будущем.

Риск — это количественная оценка потенциальных потерь и их вероятности в силу неопределённости будущих событий.

В современных финансах есть целая наука — **риск-менеджмент**, посвящённая задачам классификации, анализа, оценки и управления рисками.

Однако для подавляющего большинства приложений риск называется **волатильностью** и измеряется как среднеквадратичное отклонение доходности финансового актива.

Основы портфельного анализа

Анализ инвестиционного портфеля

Инвестиционный портфель — набор финансовых активов, который управляется как единое целое, с учётом зависимостей доходности и риска отдельных активов между собой.

Главное отличие инвестиционного портфеля от набора активов — свойства портфеля определяются не только составом и свойствами каждого отдельного актива,
но и взаимосвязью параметров активов между собой.

Диверсификация инвестиционного портфеля

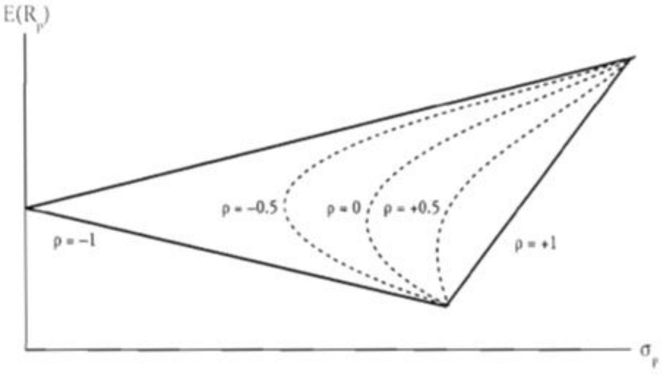
Вся современная теория управления портфелем была заложена двумя учёными-экономистами — **Гарри Марковицем** и **Уильямом Шарпом.**

Марковиц впервые математически описал эффект **диверсификации** — снижения риска инвестиционного портфеля в результате добавления в него новых активов.

Диверсификация основана на несовершенной корреляции между доходностями активов. Доходность и риск портфеля из двух активов А и В рассчитывается по формулам:

$$E(R_{portfolio}) = W_A E(R_A) + W_B E(R_B)$$

$$\sigma_{portfolio} = \sqrt{W_A^2 \sigma_A^2 + W_B^2 \sigma_B^2 + 2W_A W_B \rho_{AB} \sigma_A \sigma_B}$$



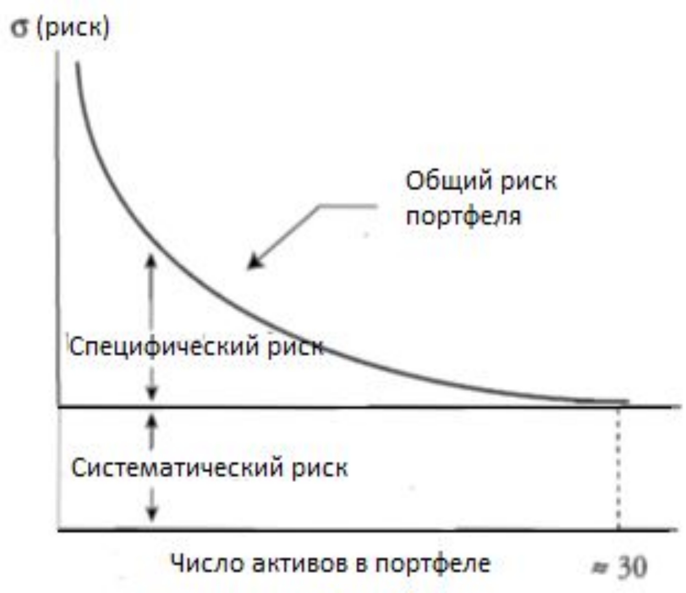
Ограничение потенциала диверсификации

Уильям Шарп впервые обратил внимание на то, что потенциал диверсификации ограничен: риск не будет падать ниже определённого уровня, который он назвал **рыночным** или **систематическим** риском.

Шарп предположил, что рынок не должен вознаграждать за тот риск, который может быть диверсифицирован, и инвестор получает премию только за **рыночный риск**.

Модель Шарпа позволяет найти равновесную (справедливую) доходность любого актива как функцию от чувствительности к рыночному риску (**коэффициента бета**):

$$E(R_i) = R_f + \beta_i[E(R_{mkt}) - R_f]$$



На уроке мы

- 📌 Познакомились с основными терминами теории вероятностей и математической статистики.
- 📌 Рассмотрели применение статистических концепций в экономике и финансах.
- 📌 Познакомились с ролью риска и неопределённости в современных финансах.
- 📌 Изучили расчёт доходности и риска финансовых активов.
- 📌 Познакомились с акциями и разобрали их роль на финансовых рынках.
- 📌 Рассмотрели основные теории и формулы портфельного анализа.
- 📌 Научились применять теорию вероятностей и математическую статистику для практического анализа инвестиционных портфелей.

Спасибо за внимание!