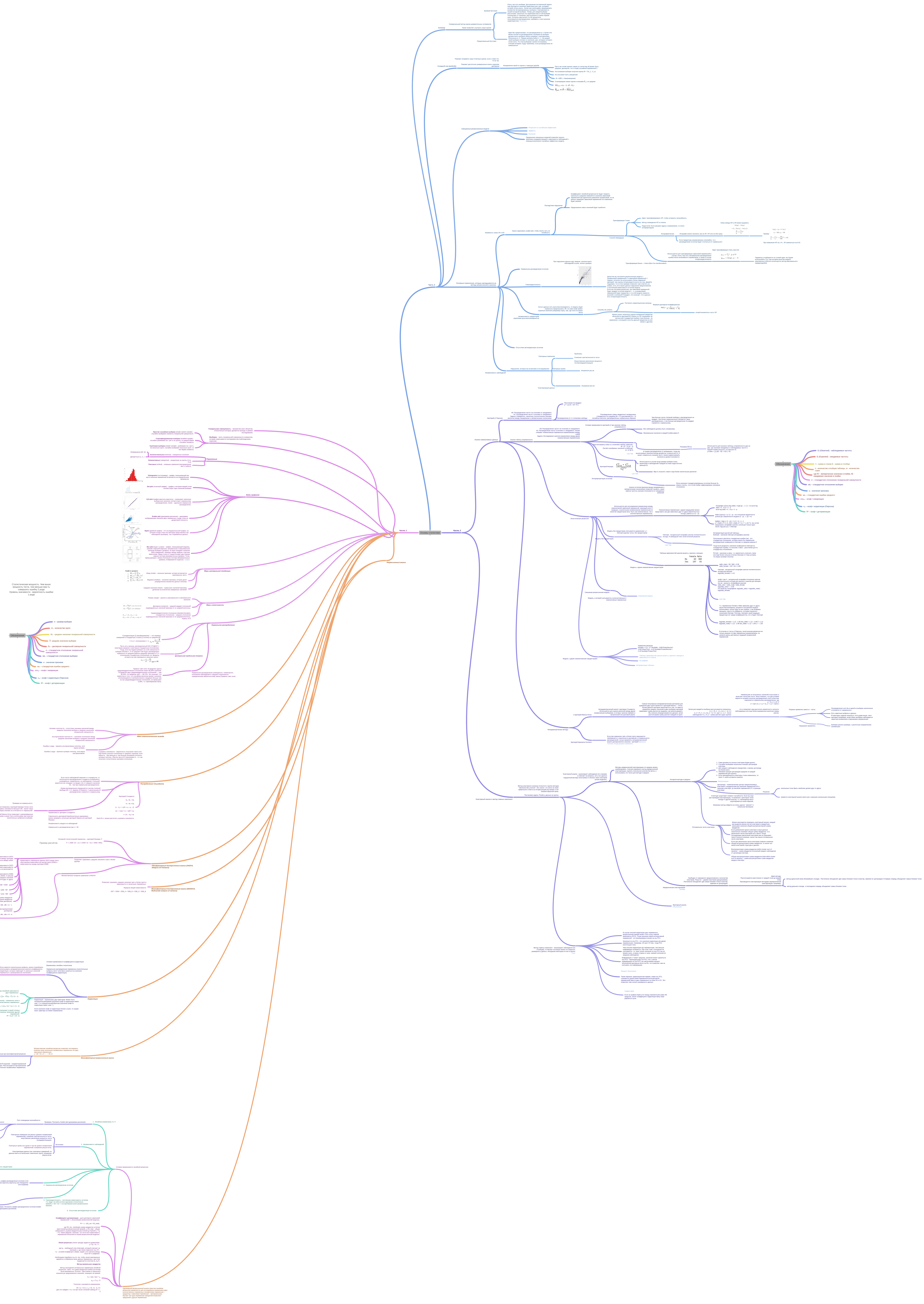
coggle
made for free at coggle.it


Основы статистики
Часть 1
Часть 2
Часть 3
Обозначения
n – размер выборки
m – количество групп
Mₓ - среднее значение генеральной совокупности
X̄- среднее значение выборки
Dₓ – дисперсия генеральной совокупности
σ – стандартное отклонение генеральной совокупности
sdₓ – стандартное отклонение выборки
xᵢ - значение признака
seₓ – стандартная ошибка среднего:
covₓᵧ – коэф-т ковариации
rₓᵧ – коэф-т корреляции (Пирсона)
R² – коэф-т детерменации
Oᵢ (Observed) - наблюдаемые частоты
Eᵢ (Expexted) - ожидаемые частоты.
fᵢ - сумма в строке fj - сумма в столбце
n - количество столбцов таблицы, m - количество строк
где fO - эмпирическое значение в ячейке, fE - ожидаемое значение в ячейке
σ – станадртное отклонение генеральной совокупности
Статистическая мощность. Чем выше мощность теста, тем меньше вер-ть совершить ошибку 2 рода
Уровень значимости - вероятность ошибки 1 рода
Генеральная совокупность – множество всех объектов, относительно которых делаются выводы в рамках исследования.
Выборка – часть генеральной совокупности элементов, которая охватывается экспериментом (наблюдением, опросом).
Переменные
Виды графиков
Histogram
Dot plot
Q-Q plot
Scatter plot
Biplot
Box plot
Меры центральной тенденции
Меры изменчивости
Нормальное распределение
Центральная предельная теорема
Идея статистического вывода
Распределение Стьюдента
Пример расчётов
Однофакторный дисперсионный анализ (ANOVA, ANalysis Of VAriance)
Многофакторный дисперсионный анализ (MANOVA, Multivariate analysis of variance)
Корреляция
Многофакторный регрессионный анализ
Условия применимости линейной регрессии:
Регрессионный анализ
Анализ номинативных данных
Критерий χ²-Пирсона
Анализ таблиц сопряжённости
Критерий Фишера
Логистическая регрессия
Непараметрические методы
U-критерий Манна-Уитни
Критерий Краскела-Уоллиса
Кластерный анализ и метод главных компонент
bootstrap
Базовый бутстрап
Перцентильный бутстрап
Складной нож (jackknife)
Смешанные регрессионные модели
Основные ограничения, которые накладываются на методы регрессионного анализа:
Линейность связи ЗП и НП
Нормальное распределение остатков
Гомоскедастичность
Независимость предикторов (проблема мультиколлинеарности)
Отсутствие автокорреляции остатков
Независимость наблюдений