

Цели и методики научного исследования

Научно-исследовательский семинар

Федор Ратников

НИУ ВШЭ
лаборатория методов анализа больших данных

Немного философии

Наука

Нау́ка — деятельность, направленная на выработку и систематизацию объективных знаний о действительности.

- Эта деятельность осуществляется путём **сбора** фактов, их регулярного **обновления**, систематизации и **критического анализа**.
- На этой основе выполняется **обобщения** или **синтез новых знаний**, которые **описывают наблюдаемые** природные или общественные явления и указывают на **причинно-следственные связи**, что позволяет осуществлять **прогнозирование**.
- Те гипотезы, которые **описывают** совокупность наблюдаемых фактов и **не опровергаются** экспериментами, признаются **законами** природы или общества

Виды наук

Естественные науки изучают строение нашего Мироздания

| физика, химия, биология (включая биоинформатику), геология, астрономия, космология, ...

Общественные науки изучают свойства человеческого мозга и связанные с ним явления

| психология, социология, ...

Формальные науки изучают свойства формальных построений безотносительных к строению Мироздания

| математика, кибернетика, теория информации, ...

Какие науки являются объективными?

Виды наук

Фундаментальные науки

цель - понять суть, ответить на вопрос “почему это так?”

Прикладные (инженерные) науки

цель - практически использовать научные знания, ответить на вопрос “как это сделать?”

- › кому-то интересней практически создавать нечто новое
- › кому-то интересней разбираться в особенностях устройства окружающей действительности

К какому типу принадлежите Вы?

Мотивы научного исследования

“Одно из наиболее сильных побуждений, ведущих к <...> науке, — это желание уйти от будничной жизни с её мучительной жестокостью и безутешной пустотой <...> Эта причина толкает людей с тонкими душевными струнами от личных переживаний в мир объективного видения и понимания. ...

К этой негативной причине добавляется и позитивная. Человек стремится каким-то адекватным способом создать в себе простую и ясную картину мира для того, чтобы оторваться от мира ощущений, чтобы в известной степени заменить этот мир созданной таким образом картиной.”

А. Эйнштейн, речь к 60-ти летию Макса Планка, 1918 г.

Роль математики

В современном мире признанным языком описания научных теорий является **математический язык**

Математическая наука строится на принципиально отличных от естественных наук принципах - на формальной логике

в математике отсутствует критерий экспериментальной проверки
математические утверждения являются точными:

- › истинными
- › ложными
- › ещё недоказанными
- › недоказуемыми

Математика и другие строго-логические дисциплины являются очень специальными типами науки

Более подробно - в лекции П.Н. Пахлова в октябре

Естественные науки

Деятельность, направленная на **выработку**
и **систематизацию объективных знаний** о **действительности**

наблюдаемость
воспроизводимость
опровержимость
совместимость

Наблюдаемость

Предмет исследования и феномены с ним связанные должны быть доступны для наблюдений

строение и законы функционирования Ада или Рая не могут быть предметом научного исследования

... а чакры могут быть предметом научного исследования?

Воспроизводимость

Результаты научных исследований, проведенных в аналогичных условиях должны воспроизводить друг друга, быть объективными

результат не должен зависеть от “тайного знания”, “магического ритуала”, секретных “know how”...

Что насчет стохастических результатов, например в квантовой механике?

Что насчёт проприетарных научных исследований, секретных разработок?

Опровержимость

Научная теория должна быть опровержимой: должен существовать априорный способ проверки возможной **не**правильности научного утверждения

- если Солнце обращается вокруг Земли, то по ходу движения Солнце должно двигаться относительно звезд...
- теория загробной жизни не является опровержимой на имеющемся уровне научных знаний

Является ли опровержимым утверждение, что любую разумную функцию можно аппроксимировать нейронной сетью?

Совместимость

Научный результат должен вписываться в систему уже имеющихся научных результатов

известно большое количество установленных и перепроверенных научных результатов

- › новая научная теория должна им соответствовать
- › наука развивается путем уточнения предыдущих знаний

релятивистская механика расширяет ньютоновскую механику при скоростях, близких к скорости света

Возможны ли революционные перевороты в науке?

Невозможность подтверждения научной теории

Критерий правильности теории - непротиворечивость любым, в том числе будущим, наблюдением

невозможно проверить в настоящем
истинность теории не может быть доказана строго
индуктивно

Принцип истинности научной теории - непротиворечие существующим наблюдениям - **абдукция** (сравните с индукцией):

если оно выглядит, как утка, крякает, как утка, плавает, как утка, то это - утка



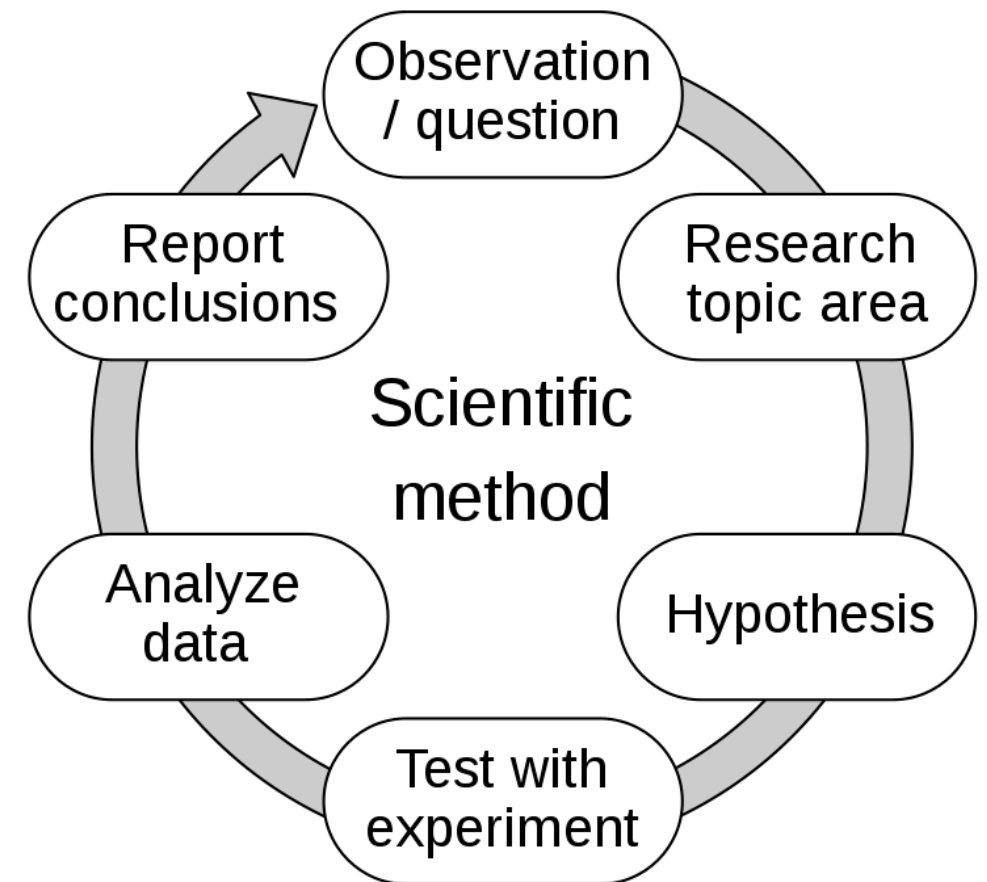
Развитие науки - научный метод

Поступательное развитие

новые исследования улучшают и/или уточняют имеющиеся знания

Цикл научного развития:

... наблюдение
осмысление
гипотеза
предсказание
экспериментальная проверка
наблюдение ...



By Efbrazil, <https://commons.wikimedia.org/w/index.php?curid=102392470>

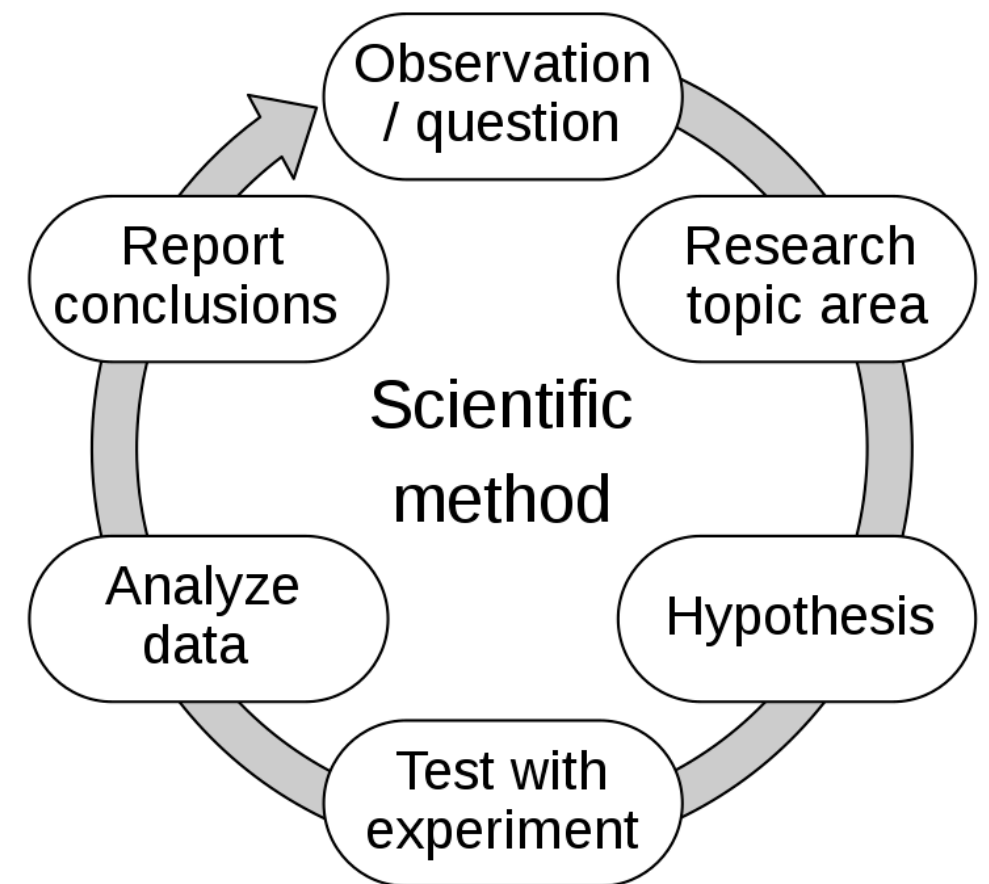
Развитие науки - научный метод

Цикл научного развития:

... наблюдение
осмысление
гипотеза
предсказание
экспериментальная проверка
наблюдение ...

В современных науках редко когда один человек или коллектив занимается всем циклом научного развития

понимание собственной роли в этом цикле
проверка, что остальные роли
выполняются адекватно



By Efbrazil, <https://commons.wikimedia.org/w/index.php?curid=102392470>

Практика

Экспериментальная проверка научной гипотезы



Гипотеза: шар упадет за 3.2 сек

Эксперимент: шар упал за 3 сек

Гипотеза подтверждена или опровергнута?

Экспериментальная проверка научной гипотезы

Гипотеза: шар упадет за 3.2 сек

Эксперимент: шар упал за 3 сек

Гипотеза подтверждена или опровергнута?

Зависит от точности измерения:

Эксперимент: шар упал за (3.00 ± 0.01) сек

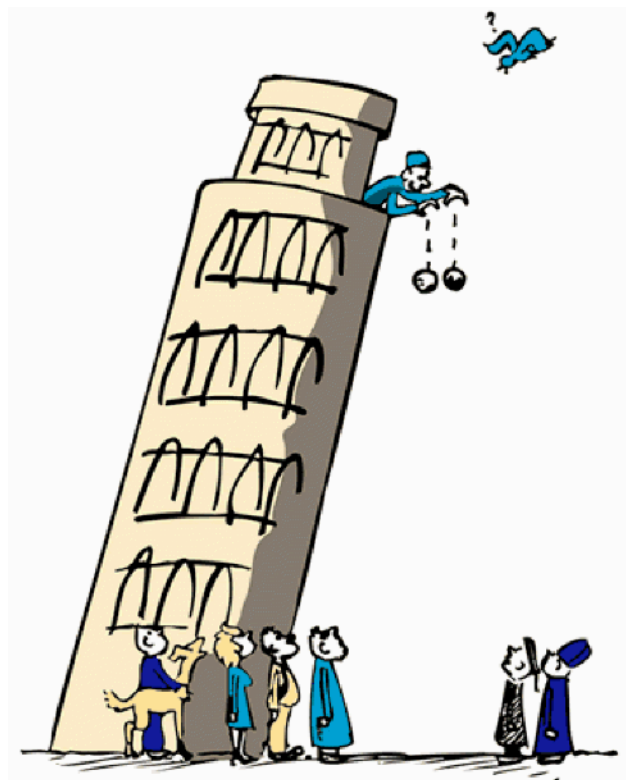
› гипотеза отвергнута (на каком CL?)

Эксперимент: шар упал за (3.0 ± 0.3) сек

› согласие с гипотезой (на каком CL?)

Эксперимент: шар упал за (3 ± 1) сек

› недостаточно точности измерения



Пример из CS

Имеем задачу классификации

имеем датасет, делим на тест и трейн

Построили два классификатора

натренировали на трейне
получили скоры на тесте

Результат

(1) ROC AUC = 0.8934578

(2) ROC AUC = 0.9015689

Какой классификатор лучше с научной точки зрения?

Неопределенность предсказаний



Гипотеза: шар упадет за 3.2 сек

Эксперимент: шар упал за (3.30 ± 0.01) сек

Гипотеза подтверждена или опровергнута?

Неопределенность предсказаний

Гипотеза: шар упадет за 3.2 сек

Эксперимент: шар упал за (3.30 ± 0.01) сек

Гипотеза подтверждена или опровергнута?

Какова точность предсказания?

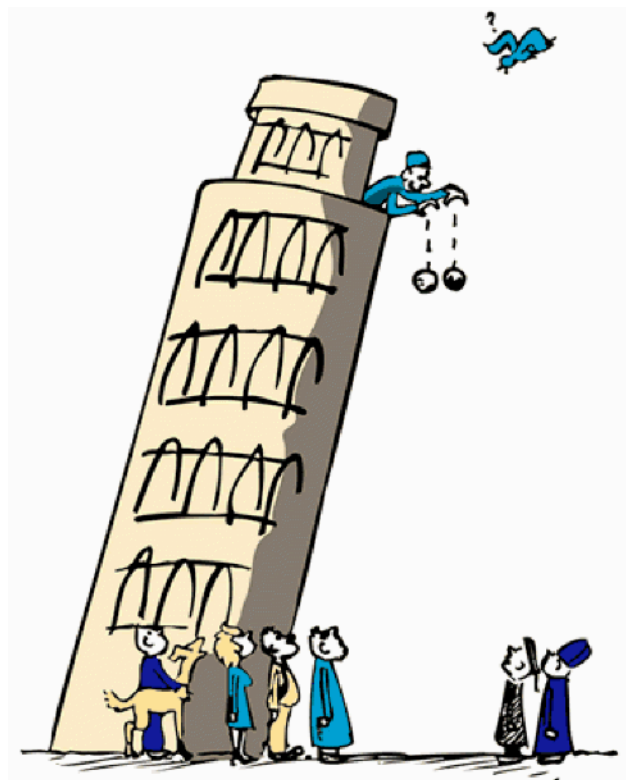
$$t = \sqrt{2H/g} ?$$

каков эффект сопротивления воздуха?

насколько точным является предсказание?

Если предсказание $(3.2^{+0.1}_{-0.0})$ сек, то согласие вполне приличное (какой CL?)

Кстати, что означает $(3.2^{+0.1}_{-0.0})$?



Точность предсказаний и измерений

Практически каждое численное научное измерение и/или предсказание сопряжено с неопределенностью в определении этого числа

часто называют “ошибкой”, “погрешностью”, но без отрицательной коннотации

Научная интерпретация чисел возможна только в совокупности с оценками их неопределенностей

численные результаты без явных оценок неопределенностей редко имеют научный смысл

Пример из CS

“Мы показали, что алгоритм А превосходит алгоритм Б на изученных датасетах”

Насколько это утверждение является научным?

Статистика

В чем разница между статистикой и теорией вероятности?

Частотный подход к вероятности

Вероятность исхода - предел относительного количества данных исходов в пределе большого количества экспериментов в идентичных условиях

где взять бесконечные эксперименты?

как быть с единичными событиями

› $p(\text{завтра дождь}) = ?$

Субъективный байесовский подход

Вероятность исхода - степень уверенности (персональной, субъективной)

| готов поставить 5:1, что на кубике не выпадет “6”

› $p(\text{not “6”}) = 5/6$

| готов поставить 2:1, что завтра будет дождь

› $p(\text{завтра дождь}) = 2/3$

Теорема Байеса позволяет манипулировать условными вероятностями

|
$$p(A | B) = \frac{p(B | A)p(A)}{p(B)}$$

› в частности вычислять апостериорную вероятность на основе экспериментальных наблюдений

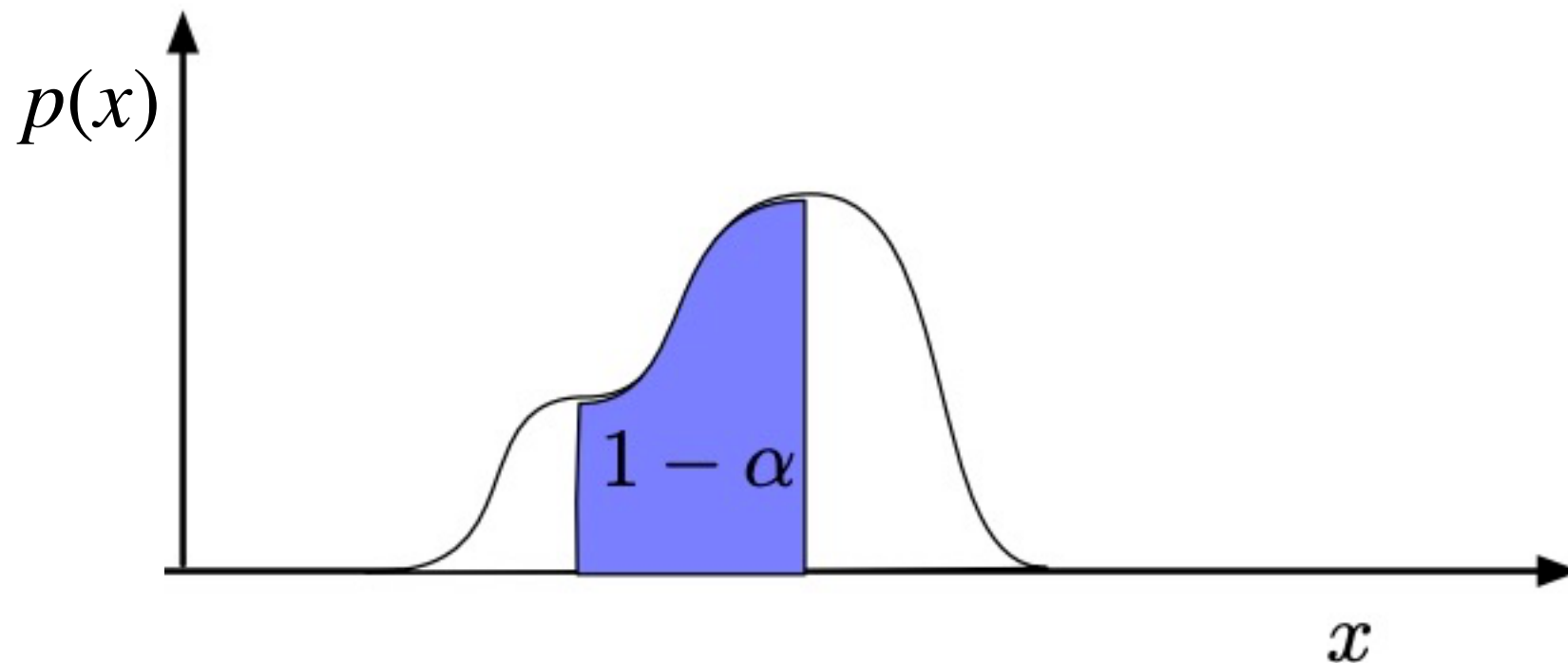
Можно ли использовать субъективную байесовскую вероятность для научных измерений?

Доверительный интервал

Имеем распределение некоторой случайной величины

хотим найти интервал в котором лежит большая часть распределения

$$\rangle p(x \in I) = 1 - \alpha$$



Насколько однозначное определение?

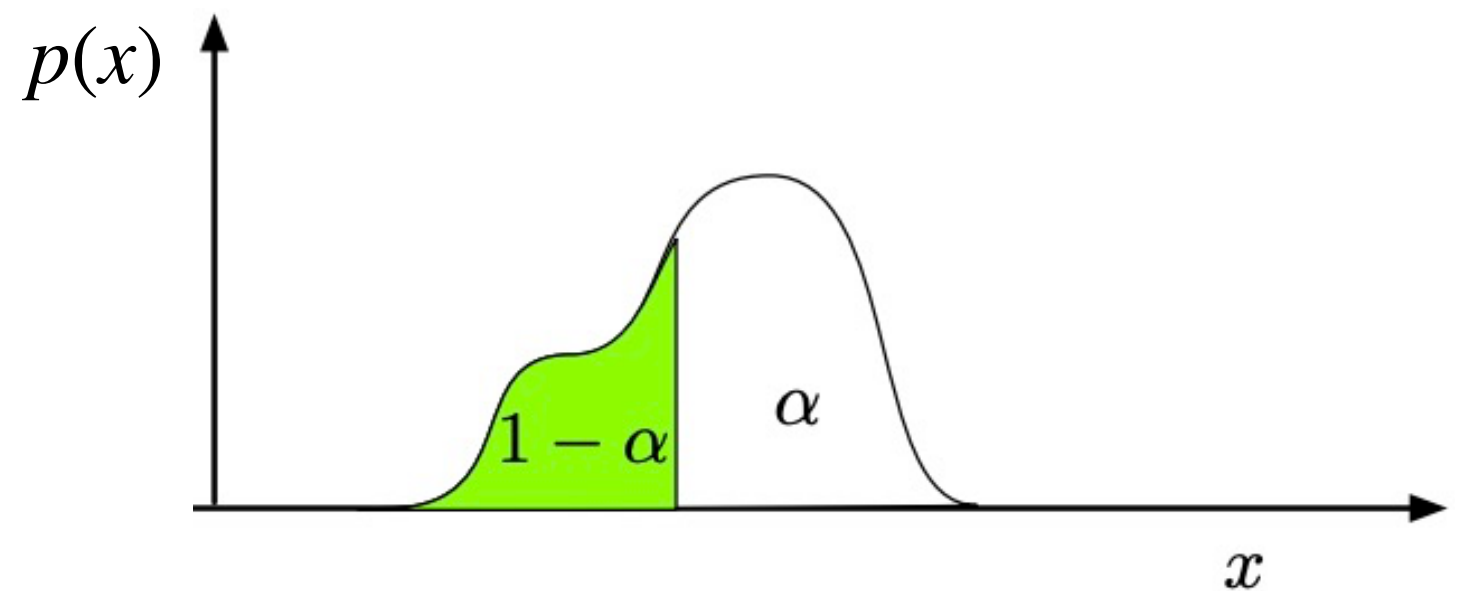
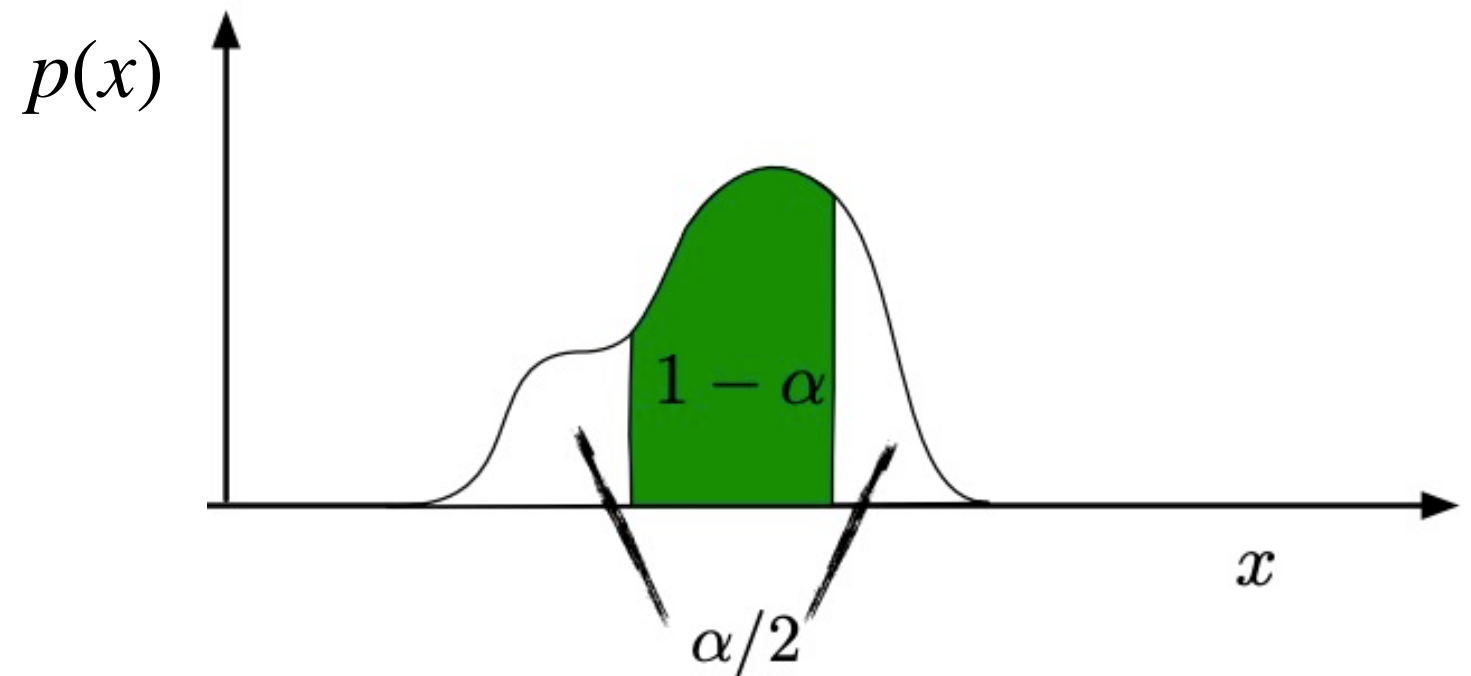
Доверительный интервал

■ центральный интервал

■ односторонний интервал

■ минимальный интервал

...



Результат научного исследования

Обычно выражается в утверждении типа:

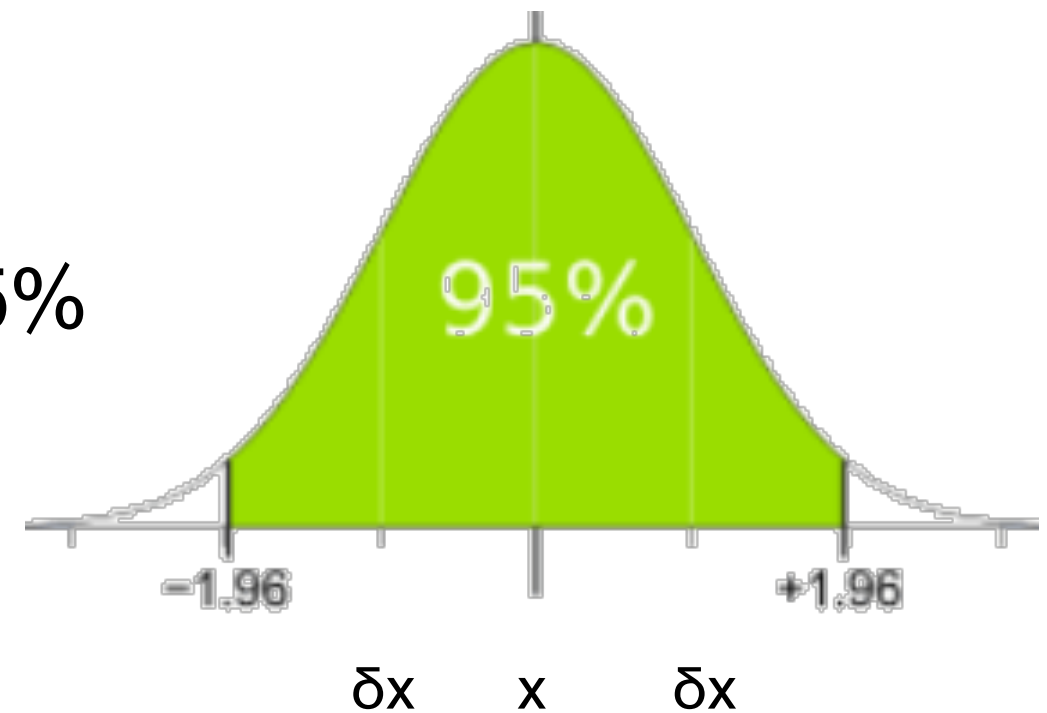
“В результате исследования значение параметра θ с 90% вероятностью лежит в интервале $a < \theta < b$ ”?

Что означает это выражение?

Это то же самое, что $\theta = \frac{a+b}{2} \pm \frac{b-a}{2}$?

Confidence Interval

- ◇ Say 95% CL
 - ◇ make many observations
 - ◇ select (central) interval containing 95% of observations
 - ◇ $x_0 \pm \delta x$
- ◇ then for the observation x 95% confidence interval is $x \pm \delta x$



Confidence Interval

- ◇ Say 95% CL
- ◇ make many observations
- ◇ select (central) interval containing 95% of observations
- ◇ $x_0 \pm \delta x$
- ◇ then for the observation x 95% confidence interval is $x \pm \delta x$



Определение доверительного интервала

У нас есть некоторый измерительный инструмент

■ для каждого истинного значения измеряемого параметра θ инструмент выдает наблюдаемое значение x с некоторым стохастическим распределением $p(x | \theta)$

› кстати, откуда мы знаем $p(x | \theta)$?

калибровка инструмента (обмер эталона)

■ мы сделали однократное измерение априори неизвестного параметра θ^* и получили значение x^*

Как определить доверительный интервал для величины θ^* с заданным покрытием $1 - \alpha$?

Должны ли измеряемое θ и наблюдаемое x иметь одинаковые размерности?

В байесовском случае всё просто

Интервал - заданный вероятностный диапазон значений параметра

$$I : p(\theta \in I) = \int_I d\theta p(\theta | x) = \int_I d\theta \frac{p(x | \theta)\pi(\theta)}{\int_I d\theta p(x | \theta)\pi(\theta)} = 1 - \alpha$$

достоверный интервал a.k.a. credible interval

В частотном случае чуть сложнее

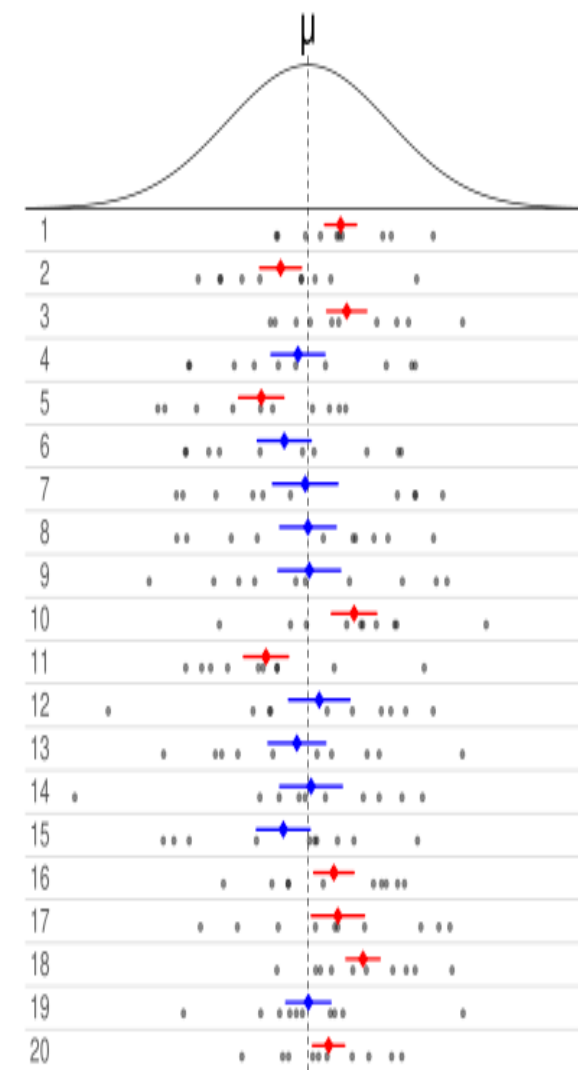
доверительный интервал a.k.a. confidence interval

у нас есть истинное значение θ_0
мы делаем много измерений, для каждого измерения
определяем доверительный интервал согласно
некоторой процедуре

› эти интервалы различны ввиду стохастичности
измерений

истинное значение θ_0 находится внутри ровно в $1 - \alpha$
части случаев

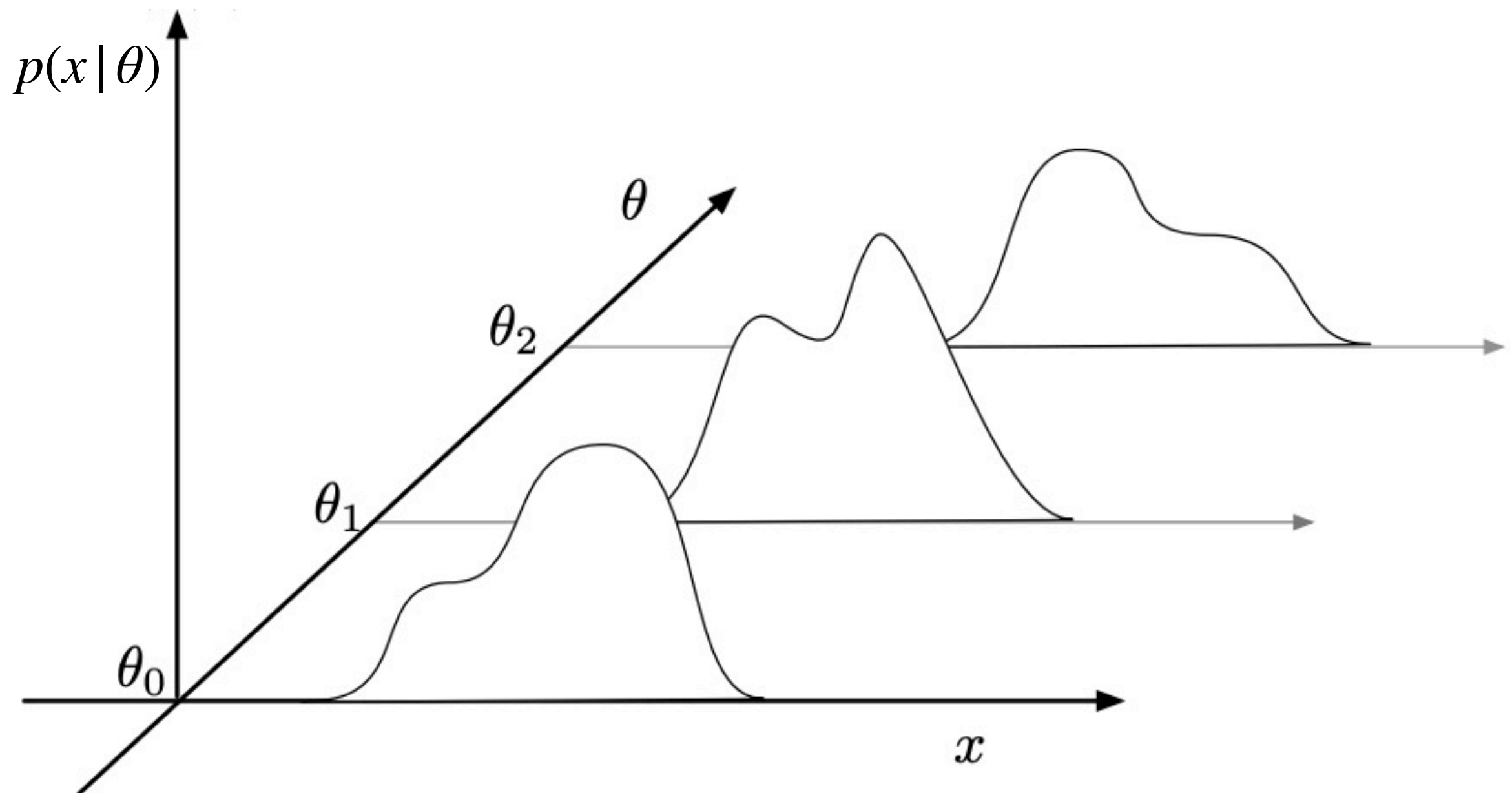
› требование правильного “покрытия” доверительного
интервала



Как построить такой интервал?

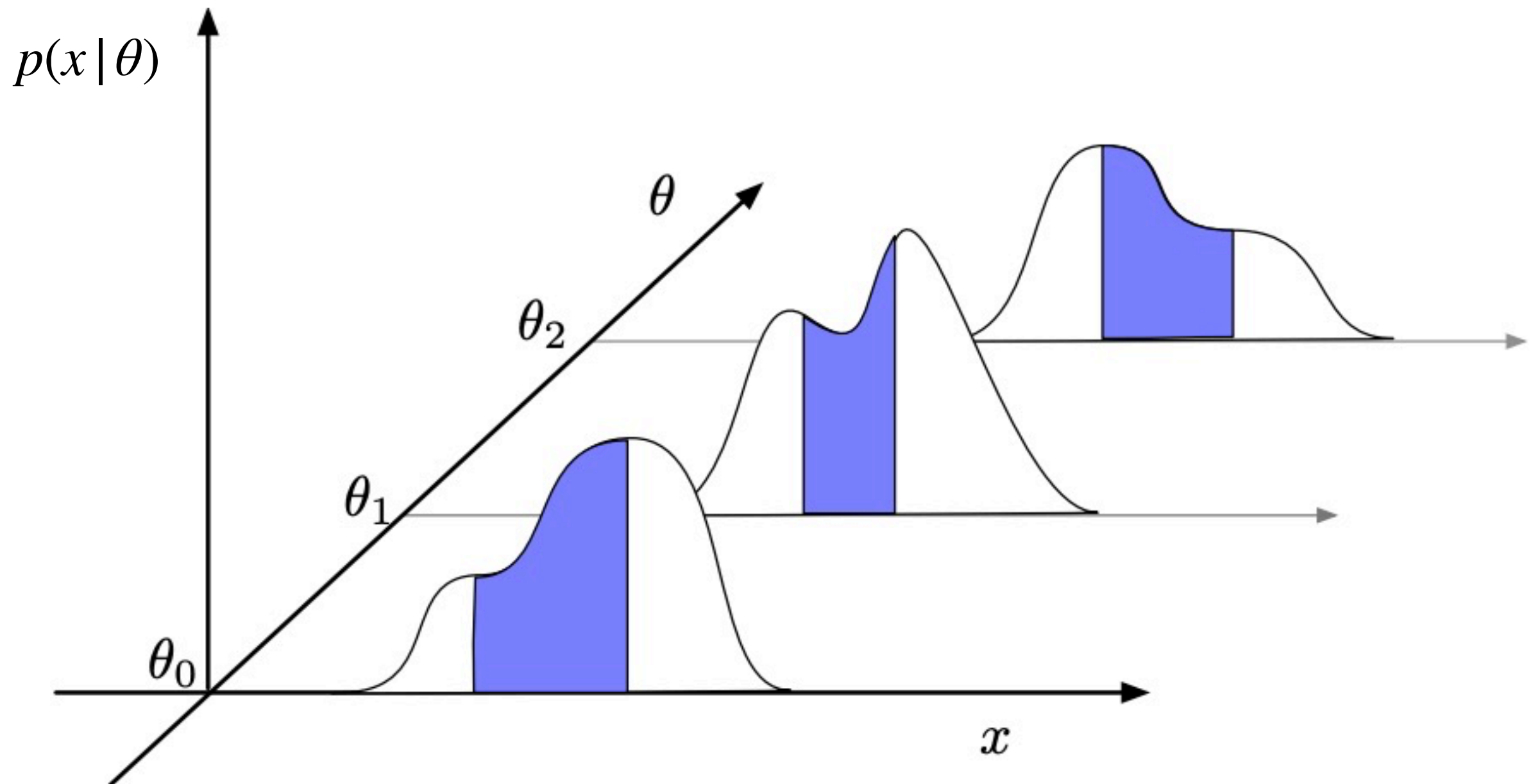
Neyman construction, шаг 1

Для каждого измеряемого θ определяем распределение наблюдаемых $p(x | \theta)$



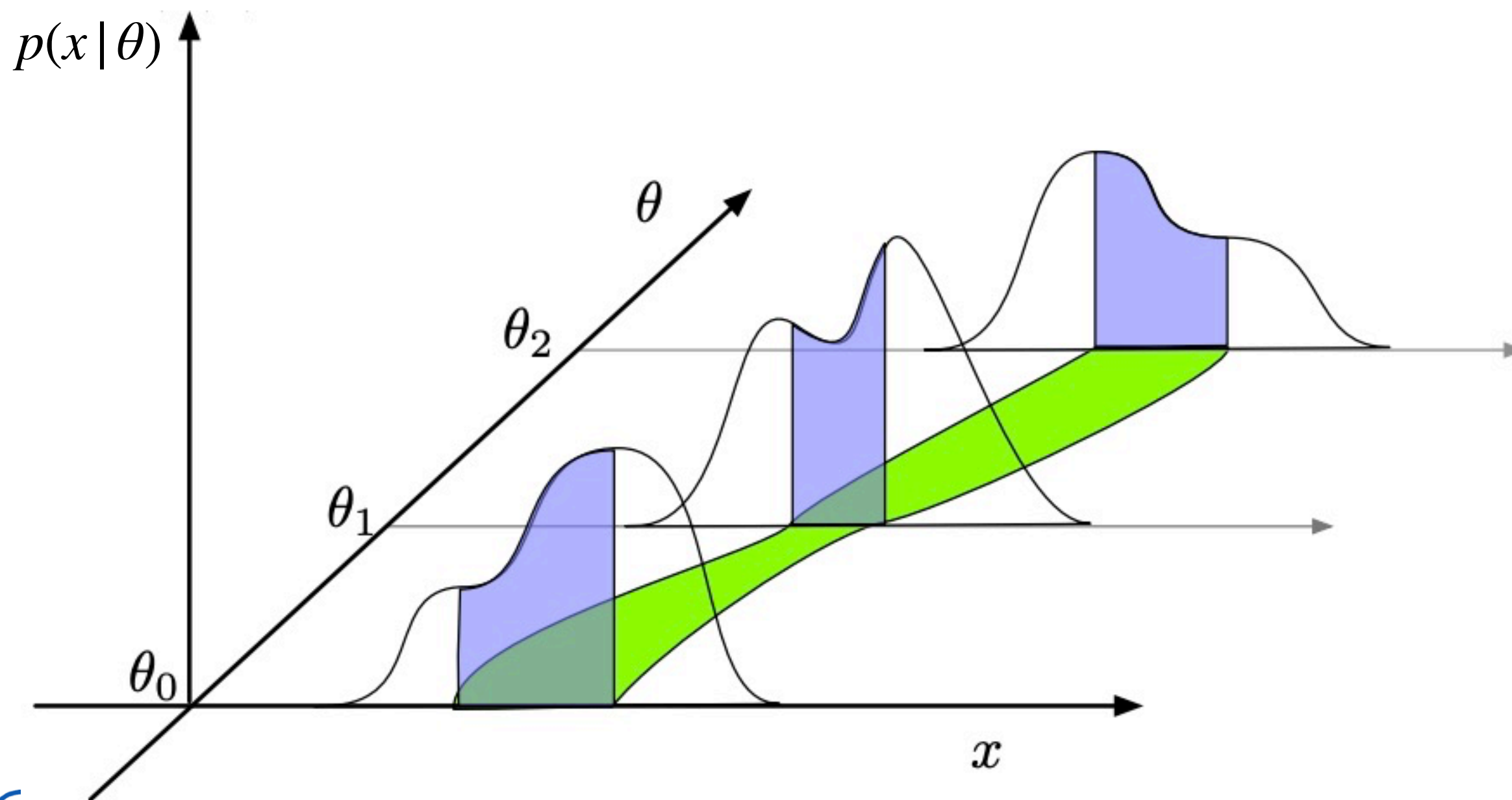
Neyman construction, шаг 2

Для каждого распределения наблюдаемых $p(x | \theta)$ определяем интервал, соответствующий $1 - \alpha$



Neyman construction, шаг 3

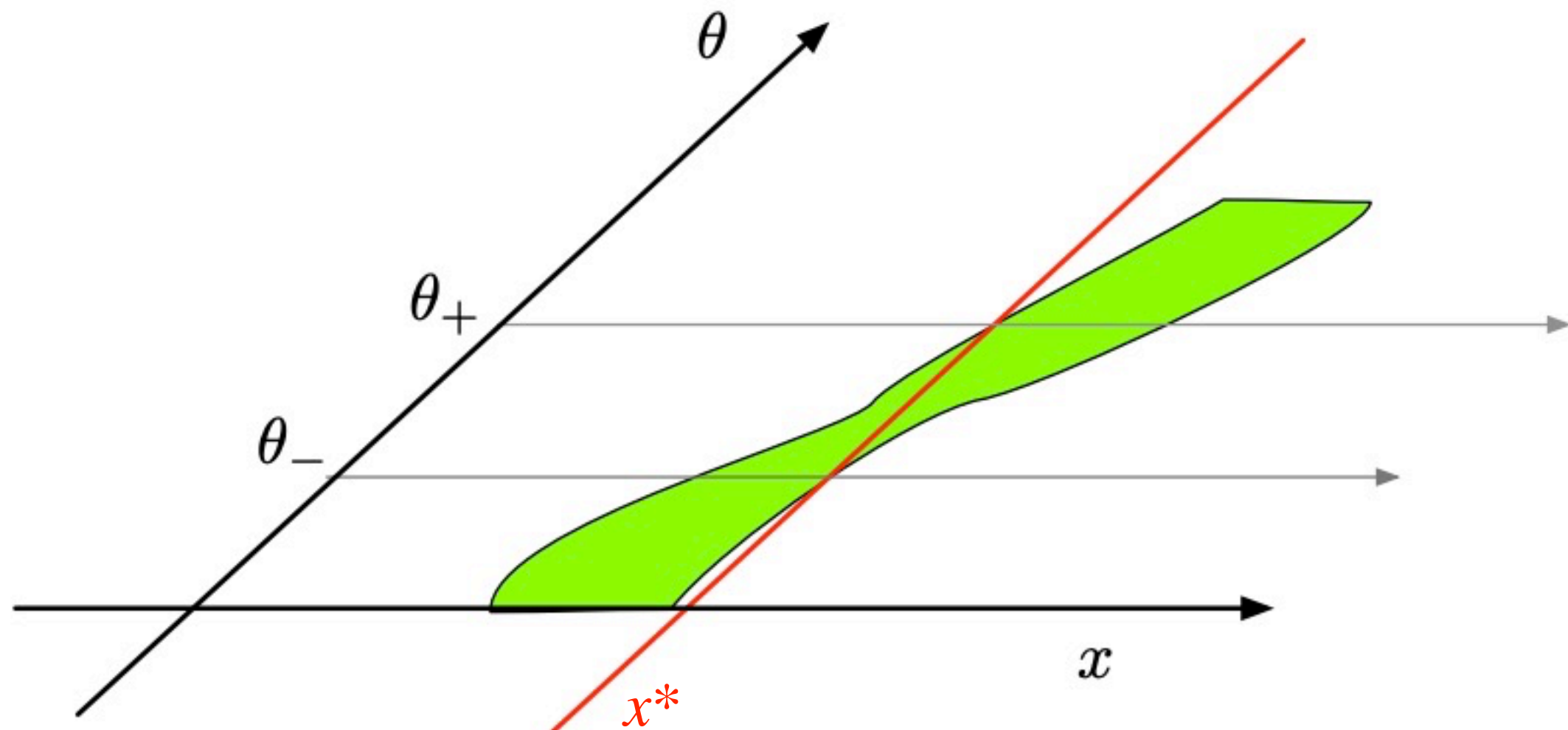
Получаем соответствующий домен в плоскости
наблюдаемые:измеряемые (x, θ)



Neyman construction, шаг 4

Накладываем наше измерение x^*

- интервал пересечения с доменом задает искомый доверительный интервал $\theta^* \in [\theta_-, \theta_+]$ с уровнем достоверности $1 - \alpha$



Частотный или байесовский интервал

Частотный подход

- хорошо определенное объективное определение частотной вероятности

- построение интервалов требует некоторой аккуратности
- полученный интервал непросто использовать как вход для последующих измерений

Байесовский подход

- зависимость результата от субъективного априорного распределения

 - › нарушение требования воспроизводимости

- ресурсоемкие вычисления нормализующих интегралов
- полученный постериор можно использовать как вход для последующих измерений

Частотный или байесовский подход

Частотная и байесовская интерпретации измерений дают различные результаты

| который правильный?

Байесовская интерпретация предполагает, что у нас возможны различные параметров Мироздания с различными, иногда субъективными вероятностями

Естественнонаучный подход предполагает уникальность и универсальность набора параметров Мироздания, их объективную сущность

| естественное описание частотным подходом научных исследований

Практически, при большом количестве измерений, результаты байесовских вычислений становятся менее зависимыми от субъективных априорных предположений, в пределе сходясь с частотным подходом при существенно более простом математическом аппарате

Больше деталей в лекции Д.А. Деркача

Насколько научным является утверждение

“наше исследование показало, что вероятность дождя завтра - 67%”?