Probability and statistics

Alexander Khanov

PHYS6260: Experimental Methods is HEP
Oklahoma State University

September 13, 2023

Repeatability and reproducibility

- Repeatability and reproducibility are principal features of an experiment, let's try to define them

Repeatability

The experiment is repeatable if we can do it again and get the same result

Reproducibility

The experiment is reproducible if someone can repeat it and get the same result

- Are these definitions good?

Repeatability and reproducibility

- Measurements never yield the same result due to
  - intrinsic principles of nature (quantum mechanics)
  - many minor factors affecting the measurement (random errors)
- Here are better definitions:

Repeatability

The experiment is repeatable if we can do it again and get a consistent result

Reproducibility

The experiment is reproducible if someone else can repeat it and get a consistent result

- The whole point here is the word "consistent"
  - to figure out what it means, we'll need to learn how to deal with results that change from experiment to experiment

What is probability?

- Consider an experiment that has a set of possible outcomes
  - a coin can land heads or tails (2 outcomes)
  - rolling a dice can give 6 possible outcomes (1, 2, 3, 4, 5, 6)
  - measuring a projection of electron's spin onto some axis has two possible results $\pm\frac{\hbar}{2}$
- The number of possibilities does not have to be finite
  - measuring the energy of a quantum harmonic oscillator can yield $\frac{\hbar\omega}{2}, \frac{3\hbar\omega}{2}, \frac{5\hbar\omega}{2}, \ldots$
- If we do the same measurement N times and get a certain output M times then the probability to get this output is M/N
  - in general, the M/N ratio will fluctuate from series to series
  - the law of large numbers: there exists "objective" probability P which is "close" to the experimental result M/N for "large" N
- Sum of probabilities of all possible outputs is 1:

$$\sum_i P_i = 1$$

Continuous probability

- Experimental output can be either a set of discrete values (finite or infinite) or any value withing a certain range (which can be limited or unlimited)
- There are no strong boundaries between the two possibilities:
  - the harmonic oscillator has continuous energy spectrum when treated classically, but discrete spectrum in quantum mechanics
  - in solid state theory, conduction and valence energy bands are technically made of discrete levels but their number is so large ($\sim N_A$) that in practice they are considered continuous
  - in a hydrogen atom the electron energy spectrum has both a continuous and a discrete part

Continuous probability (2)

- In continuous case, the probability of discrete outputs is replaced with a probability density function $f(x)$
  - the probability for output x to be between x_1 and x_2 is $P(x_1 < x < x_2) = \int_{x_1}^{x_2} f(x)\,dx$
- Similar to discrete probability, p.d.f. has to be normalized:

$$\int_{x_{min}}^{x_{max}} f(x)\,dx = 1$$

  - x_{min} can be $-\infty$ and/or x_{max} can be $+\infty$
- In general, if both discrete and continuous outputs are possible,

$$\sum_i P_i + \int_{x_{min}}^{x_{max}} f(x)\,dx = 1$$

Frequentist approach to probability

- We have a hypothesis (theory) that governs random outcomes of our experiment, this hypothesis may be true or false
- The experimental data has a certain probability which is defined as the limit $P = \lim_{N\to\infty} \frac{M}{N}$
  - we can't have an infinite series of experiments, so we take the observed value of $\frac{M}{N}$ as an approximation for P and assign some uncertainty to it
  - based on how significant is the difference between the predicted and observed value of P, we decide whether the original theory is valid

Frequentist:

there is one hypothesis that can be true or false
the experimental outcome has certain probability

Example 1

- Problem:
  - I have a coin that I tossed 10 times and got 6 heads and 4 tails. Can I reject the hypothesis that tails and heads are equally probable, at the 95% confidence level?
- Solution:
  - the probability to get 6 heads and 4 tails is expected to be

    $$\frac{10!}{6!\,4!}0.5^6(1-0.5)^4 = 0.21,$$

    which is more than 0.05, therefore the hypothesis can't be rejected
- Remarks:
  - if the outcome was 10 heads and 0 tails, the probability of such outcome would be $0.5^{10} = 10^{-3}$, so the hypothesis would have to be rejected
  - note that 0.05 is not the probability that our hypothesis is true, instead it is the probability that a true hypothesis is rejected due to randomness

Bayesian approach to probability

- We have data D which represents the result of an experiment
  - consider a set of hypotheses (theories) H_i which may yield this particular result with prior probabilities $P(D|H_i)$
  - the probability for a given hypothesis H_i to be true can be calculated from Bayes' theorem

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D)}, \quad P(D) = \sum_j P(D|H_j)P(H_j)$$

Bayesian:

there are many hypotheses, each has a probability to be true
the experimental outcome is fixed

Example 2

- Problem:
  - 5% of the population are sick
  - 98% of persons tested positive are sick
  - 4% of persons tested negative are sick
  - If a person is tested positive, what's the probability he's sick?
- Solution:
  - we have two hypotheses: a person is sick (H_1, $P(H_1) = 0.05$) and a person is healthy (H_2, $P(H_2) = 1 - P(H_1) = 0.95$)
  - the experimental result D is that the person is tested positive
  - the conditional probabilities for result D are $P(D|H_1) = 0.98$ and $P(D|H_2) = 0.04$
  - calculation yields $P(H_1|D) = \dfrac{0.98 \cdot 0.05}{0.98 \cdot 0.05 + 0.04 \cdot 0.95} = 0.56$
- Remarks:
  - Frequentist couldn't say anything definite in this situation, because there is not enough data (just one measurement)

Comparison of the two approaches

- Bayesian approach is particularly useful if the theory has certain parameters and we want to evaluate those parameters based on experimental results

Example 3

- Problem:
  - I have a coin with (unknown) heads probability p and tails probability $1 - p$. After I tossed it 14 times, I got 10 heads and 4 tails. What is the probability that in the next two tosses there will be two heads in a row?
- Frequentist solution:
  - the estimate for p from our experiment is $p = 10/14 = 0.714$
  - the probability to get two heads is $P(hh) = (0.714)^2 = 0.51$

Example 3 (2)

- Bayesian solution:

$$P(p|D) = \frac{P(D|p)P(p)}{P(D)}, \quad P(D|p) = \frac{14!}{10!\,4!} p^{10}(1-p)^4 = f(p, 11, 5)$$

where f is probability density function of the beta distribution:

$$f(p, a, b) = \frac{1}{B(a, b)} p^{a-1}(1-p)^{b-1}, \quad B(a, b) = \int_0^1 p^{a-1}(1-p)^{b-1}\, dp = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

- the choice of $P(p)$ (prior probability distribution) is up to us
- a convenient choice for $P(p)$ is a conjugate prior, such that $P(p)$ and $P(p|D)$ belong to the same pdf family: $P(p) = f(p, a, b)$

$$P(p|D) = \frac{f(p, 11, 5) f(p, a, b)}{P(D)} = const\; p^{10}(1-p)^4 p^{a-1}(1-p)^{b-1} =$$

$$const\; p^{10+a-1}(1-p)^{4+b-1} = f(p, 10+a, 4+b)$$

Example 3 (3)

- For a given p, the probability to get two heads in a row is $P(hh|p) = p^2$
- Given the results D of the experiment,

$$P(hh|D) = \int_0^1 P(hh|p)P(p|D)\,dp =$$
$$\frac{1}{B(10+a, 4+b)} \int_0^1 p^2 p^{10+a-1}(1-p)^{4+b-1}\,dp =$$
$$\frac{1}{B(10+a, 4+b)} \int_0^1 p^{12+a-1}(1-p)^{4+b-1}\,dp = \frac{B(12+a, 4+b)}{B(10+a, 4+b)}$$

- Choice 1: flat prior, uniform p distribution, $a = b = 1$, $P(p) = 1$
  - $P(hh|D) = B(13, 5)/B(11, 5) = 0.485$
- Choice 2: any mean of the p distribution is equally likely, $a = b = 0$
  - $P(hh|D) = B(12, 4)/B(10, 4) = 0.524$
- In general, if in a preliminary series we have seen a heads and b tails, then it makes sense to pick $B(a, b)$ as a prior