Probability Theory v1.0.1_20240904

- *The principle of indifference in classical probability, such as when guessing a password*

```
if (nchar(input_password) != nchar(password)) print("Try Constant-Time Crypto")
```

- *Aleatory or Epistemic. Frequentist or Bayesian.*

Probability is the long-run frequency of repeated events. $f(Y \mid \theta)$ is a function of Y with θ fixed but unknown.
Probability is a degree of uncertainty about values. $f(Y \mid \theta)$ is a function of θ with Y fixed.

- *Complete probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *and probability axioms* (Kolmogorov, 1933)

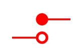
sample space $\{\omega_1, \cdots, \cancel{\omega_n}\}$, σ-algebra of subsets of Ω and event space $\cancel{2^\Omega}$, and probability measure $\mathbb{P}(E) \cancel{=} |E| \,/\, |\Omega|$

i) $\mathbb{P}(E) \geqslant 0, \forall E \in \mathscr{F}$ ii) $\mathbb{P}(\Omega) := 1$ iii) $\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$, where $E_i \cap E_j = \varnothing$ for $i \neq j$

- *Neither "random" nor "variable". Which expression is more rigorous in its formulation?*

$X(\omega) = s,\ s \in \{1,2,3,4,5,6\} \subset \mathbb{N}^+$ or $X(\omega) = i \cdot \mathbf{1}_{S_i}(\omega),\ i \in \{1,2,3,4,5,6\}$

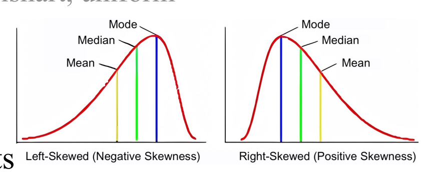
random variable and vector - measurable function $X : \Omega \to R$ measurable space, usually $R = \mathbb{R}^n$

- `geom_step()` *is a non-decreasing and right-continuous function.* 

cumulative **d**istribution **f**unction $F_X(x) := \mathbb{P}(X \leqslant x)$, e.g., the survival function $S(t) := \mathbb{P}(T > t) = 1 - F_T(t)$

- *Can a "zero-probability" event happen?* $\mathbb{P}(X = x) = 0, \forall x \in \mathbb{R}$ *but* $\mathbb{P}(X \in \mathbb{R}) = 1$ *almost surely.*

discrete RV and **p**robability **m**ass function: Bernoulli, binomial, Poisson, hyper geometric, negative binomial, uniform
continuous RV and **p**robability **d**ensity function: normal, exponential, Cauchy, triangular, gamma, beta, Laplace, Weibull, Pareto, Pearson, Dirichlet, Wishart, uniform
sampling distribution: χ^2, F, t ; quantile function, p-value



- *Which distributed RV has an undefined expected value?* $\mathrm{Kurt}(X) \neq 3 \implies$ non-normal

mean, median, and mode, variance and covariance, skewness, kurtosis, and moments
moment generating function, characteristic function, Fisher information, entropy, statistical distance

- *Being **marginally** normally distributed and linearly uncorrelated does not necessarily imply independence!*

- *Note P and p.*

$$\underbrace{\frac{P(\mathscr{M}_1 \mid d)}{P(\mathscr{M}_0 \mid d)}}_{\text{posterior odds}} = \underbrace{\frac{p(d \mid \mathscr{M}_1)}{p(d \mid \mathscr{M}_0)}}_{\text{Bayes factor } BF_{10}} \cdot \underbrace{\frac{P(\mathscr{M}_1)}{P(\mathscr{M}_0)}}_{\text{prior odds}}$$

(Ly et al., 2016, p. 21)

We use the term probability loosely in the development to refer either to probability mass or to probability density, depending on whether the data are discrete or continuous (Rouder et al., 2012, p. 358-359).

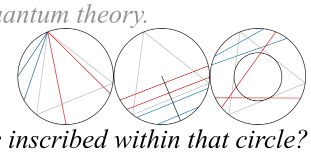
Bayesian one-sample t-test:

$$JZSBF_{10}\left(t(n),\, n\right) = (2\pi)^{-\frac{1}{2}} h \left(1 + \frac{t^2}{n-1}\right)^{\frac{n}{2}} \int_0^{\infty} (1+ng)^{-\frac{1}{2}} \left(1 + \frac{t^2}{(1+ng)(n-1)}\right)^{-\frac{n}{2}} g^{-\frac{3}{2}} e^{-\frac{h^2}{2g}}\, \mathrm{d}g$$

Discussion

1. Are these related to random events? A collection of subsets of the sample space that is closed under complementation and countable unions.

   - *The smallest σ-field of subsets of $\Omega = \{a, b, c, d\}$ that contains $\{a\}$.* *Rubik's cube, Graham Parker 26 years,* $|\Omega| \approx 4.3 \times 10^{19}$
   - *Two positive integers chosen at "random" are co-prime.* $6/\pi^2$
   - *The Fermi paradox in the search for extraterrestrial life.* $N = R_* \cdot f_p \cdot n_e \cdot f_l \cdot f_i \cdot f_c \cdot L$
   - *How many golfballs can fit in a Boeing 747?*
   - *Steve, who is an introvert and likes reading, is a librarian / a farmer.*
   - *If the forecaster was 80% certain that rain would develop but only expected to cover 50% of the forecast area, then the forecast would read "a 40% chance of rain" for any given location.*
   - *A sun would rise within a three-body system. Chaos theory.*
   - *An atomic orbital describes the location and wave-like behavior of an electron in an atom. Quantum theory.*



2. The Bertrand paradox and the maximum ignorance principle

   What is the probability that a chord of a circle is longer than the side of an equilateral triangle inscribed within that circle?

3. The "elevator" paradox (Wei, 2011)

   What is the expected number of stops an elevator will make when each of the m people is going to any of the n floors?

$m = n = 4$

$X =$	1	2	3	4
$\mathbb{P}(X = x) =$	$\binom{4}{1} \Big/ 4^4$	$\left(\binom{4}{1} + \binom{4}{2}/2!\right) \cdot \binom{4}{2} 2! \Big/ 4^4$	$\left(\binom{4}{2}\binom{2}{1}/2!\right) \cdot \binom{4}{3} 3! \Big/ 4^4$	$4!\ /\ 4^4$

$\mathbb{E}[X] = n(1 - (1 - 1/n)^m)$ but $X = 0$? $\lim_{n \to \infty} \mathbb{E}[X] = m$

4. Ask ChatGPT 4o about expected value (wrong)

   Would you gamble on a 50% chance of an asset increasing by 80% and a 50% chance of it decreasing by 50%?

5. Not **independent**. From $0.5^{100} \approx 0.0000000000000000000000000000789\,\%$ to $1 - \sum_{k=51}^{100} 1/k \approx 30.7\,\%$ (Miltersen, 2003)

   100 prisoners, numbered 1 to 100. 100 boxes, each containing a unique number from 1 to 100, randomly assigned.
   Each prisoner can open up to 50 boxes. The goal is for each prisoner to find their own number.
   If all prisoners succeed, they are freed; if any fail, all are executed.

6. When Zeno's *new* tortoise meets **convergence**

   Imagine a tortoise running in a straight line for an infinite time, with Achilles starting 1 meter behind, both moving in the same direction. If Achilles is always faster, will he eventually catch up to the tortoise? Suppose the speed difference is e^{-t}.

7. Let's explore large samples. Buffon's needle experiment and the Monte Carlo method.

   Suppose we have a floor made of parallel strips of wood, each the same width t, and we drop a needle with a length l onto the floor. The probability that the needle will lie across a line between two strips is $\frac{2l}{\pi t}$. *Hence,* $\hat{\pi} = \frac{2l}{t} \cdot \frac{\textit{number of trials}}{\textit{number of hits}}$.

   Lazzerini (1901) claimed 1,808 hits out of 3,408 trials ($\hat{\pi} \approx 3.1415929$). Is this possible? $|\hat{\pi} - \pi| = O(1/\sqrt{n})$

Hall, W. J., & Oakes, D. (2024). *A Course in the Large Sample Theory of Statistical Inference*. Boca Raton: CRC Press.
Kolmogorov, A. N. (1956). *Foundations of the Theory of Probability* (2nd ed.). New York: Chelsea.
Ly, A., Verhagen, J., & Wagenmakers, E.-J. (2016). Harold Jeffreys's default Bayes factor hypothesis tests: Explanation, extension, and application in psychology. *Journal of Mathematical Psychology*, *72*, 19–32. DOI: 10.1016/j.jmp.2015.06.004
Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*, 356–374. DOI: 10.1016/j.jmp.2012.08.001

U.S. National Weather Service. https://www.weather.gov/media/pah/WeatherEducation/pop.pdf
Veritasium (2022). The riddle that seems impossible even if you know the answer. https://youtu.be/iSNsgj1OCLA
Wei, Z. (2011). "电梯概率"导出的悖论与猜想. 大学生科技创新与社会实践论文集.
https://en.wikipedia.org/wiki/Misconceptions_about_the_normal_distribution https://en.wikipedia.org/wiki/Buffon%27s_needle_problem
https://www.bilibili.com/video/BV1Ay421v7J2/ https://www.bilibili.com/video/BV1hw411U7YP/ https://www.bilibili.com/video/BV1qiv1ezE9N/