WIKIPEDIA

# Poisson binomial distribution

In probability theory and statistics, the **Poisson binomial distribution** is the discrete probability distribution of a sum of independent Bernoulli trials that are not necessarily identically distributed. The concept is named after Siméon Denis Poisson.

In other words, it is the probability distribution of the number of successes in a sequence of *n* independent yes/no experiments with success probabilities $p_1, p_2, \dots, p_n$. The ordinary binomial distribution is a special case of the Poisson binomial distribution, when all success probabilities are the same, that is $p_1 = p_2 = \cdots = p_n$.

| Poisson binomial | |
|---|---|
| **Parameters** | $\mathbf{p} \in [0,1]^n$ — success probabilities for each of the *n* trials |
| **Support** | $k \in \{0, \dots, n\}$ |
| **pmf** | $\sum\limits_{A\in F_k} \prod\limits_{i\in A} p_i \prod\limits_{j\in A^c} (1-p_j)$ |
| **CDF** | $\sum\limits_{l=0}^{k} \sum\limits_{A\in F_l} \prod\limits_{i\in A} p_i \prod\limits_{j\in A^c} (1-p_j)$ |
| **Mean** | $\sum\limits_{i=1}^{n} p_i$ |
| **Variance** | $\sigma^2 = \sum\limits_{i=1}^{n} (1-p_i)p_i$ |
| **Skewness** | $\frac{1}{\sigma^3} \sum\limits_{i=1}^{n} (1-2p_i)(1-p_i)p_i$ |
| **Ex. kurtosis** | $\frac{1}{\sigma^4} \sum\limits_{i=1}^{n} (1-6(1-p_i)p_i)(1-p_i)p_i$ |
| **MGF** | $\prod\limits_{j=1}^{n} (1-p_j+p_j e^t)$ |
| **CF** | $\prod\limits_{j=1}^{n} (1-p_j+p_j e^{it})$ |

## Contents


Mean and variance
Probability mass function
Entropy
Chernoff bound
See also
References


## Mean and variance

Since a Poisson binomial distributed variable is a sum of *n* independent Bernoulli distributed variables, its mean and variance will simply be sums of the mean and variance of the *n* Bernoulli distributions:

$$\mu = \sum_{i=1}^{n} p_i$$

$$\sigma^2 = \sum_{i=1}^{n} (1-p_i)p_i$$

For fixed values of the mean ($\mu$) and size (*n*), the variance is maximal when all success probabilities are equal and we have a binomial distribution. When the mean is fixed, the variance is bounded from above by the variance of the Poisson distribution with the same mean which is attained asymptotically as *n* tends to infinity.

## Probability mass function

The probability of having *k* successful trials out of a total of *n* can be written as the sum [1]

$$\Pr(K=k) = \sum_{A\in F_k} \prod_{i\in A} p_i \prod_{j\in A^c} (1-p_j)$$

where $F_k$ is the set of all subsets of *k* integers that can be selected from {1,2,3,...,*n*}. For example, if *n* = 3, then $F_2 = \{\{1,2\},\{1,3\},\{2,3\}\}$. $A^c$ is the complement of $A$, i.e. $A^c = \{1,2,3,\dots,n\} \setminus A$.

$F_k$ will contain $n!/((n-k)!k!)$ elements, the sum over which is infeasible to compute in practice unless the number of trials *n* is small (e.g. if *n* = 30, $F_{15}$ contains over $10^{20}$ elements). However, there are other, more efficient ways to calculate $\Pr(K=k)$.

As long as none of the success probabilities are equal to one, one can calculate the probability of *k* successes using the recursive formula [2] [3]

$$\Pr(K=k) = \begin{cases} \prod\limits_{i=1}^{n} (1-p_i) & k=0 \\ \frac{1}{k} \sum\limits_{i=1}^{k} (-1)^{i-1} \Pr(K=k-i) T(i) & k>0 \end{cases}$$

where

$$T(i) = \sum_{j=1}^{n} \left(\frac{p_j}{1-p_j}\right)^i.$$

The recursive formula is not numerically stable, and should be avoided if $n$ is greater than approximately 20. Another possibility is using the discrete Fourier transform .[4]

$$\Pr(K=k) = \frac{1}{n+1} \sum_{l=0}^{n} C^{-lk} \prod_{m=1}^{n} \left(1+(C^l-1)p_m\right)$$

where $C = \exp\left(\frac{2i\pi}{n+1}\right)$ and $i = \sqrt{-1}$.

Still other methods are described in [5] .

# Entropy

There is no simple formula for the entropy of a Poisson binomial distribution, but the entropy is bounded above by the entropy of a binomial distribution with the same number parameter and the same mean. Therefore, the entropy is also bounded above by the entropy of a Poisson distribution with the same mean. [6]

The Shepp–Olkin conjecture, due to Lawrence Shepp and Ingram Olkin in 1981, states that the entropy of a Poisson binomial distribution is a concave function of the success probabilities $p_1, p_2, \dots, p_n$.[7] This conjecture was proved by Erwan Hillion and Oliver Johnson in 2015.[8]

# Chernoff bound

The probability that a Poisson binomial distribution gets large, can be bounded using its moment generating function:

$$\begin{aligned}\Pr[S \ge s] &\le E[\exp[t\sum_i X_i]]\exp(-st)\\ &= \exp\left(\sum_i 1 + (e^t - 1)p_i\right)\exp(-st)\\ &\le \exp\left(\sum_i \exp((e^t-1)p_i) - st\right)\\ &= \exp\left(s - \mu - s\log\frac{s}{\mu}\right)\end{aligned},$$

where we took $t = \log(s/\sum_i p_i)$. This is similar to the tail bounds of a binomial distribution.

# See also

- Le Cam's theorem

# References


1. Wang, Y. H. (1993). "On the number of successes in independent trials" (http://www3.stat.sinica.edu.tw/statistica/oldpdf/A3n23.pdf) (PDF). *Statistica Sinica*. **3** (2): 295–312.
2. Shah, B. K. (1994). "On the distribution of the sum of independent integer valued random variables". *American Statistician*. **27** (3): 123–124. JSTOR 2683639 (https://www.jstor.org/stable/2683639).
3. Chen, X. H.; A. P. Dempster; J. S. Liu (1994). "Weighted finite population sampling to maximize entropy" (http://www.people.fas.harvard.edu/~junliu/TechRept/94folder/cdl94.pdf) (PDF). *Biometrika*. **81** (3): 457. doi:10.1093/biomet/81.3.457 (https://doi.org/10.1093%2Fbiomet%2F81.3.457).
4. Fernandez, M.; S. Williams (2010). "Closed-Form Expression for the Poisson-Binomial Probability Density Function". *IEEE Transactions on Aerospace and Electronic Systems*. **46** (2): 803–817. Bibcode:2010ITAES..46..803F (http://adsabs.harvard.edu/abs/2010ITAES..46..803F). doi:10.1109/TAES.2010.5461658 (https://doi.org/10.1109%2FTAES.2010.5461658).
5. Chen, S. X.; J. S. Liu (1997). "Statistical Applications of the Poisson-Binomial and conditional Bernoulli distributions" (http://www3.stat.sinica.edu.tw/statistica/password.asp?vol=7&num=4&art=4). *Statistica Sinica*. **7**: 875–892.
6. Harremoës, P. (2001). "Binomial and Poisson distributions as maximum entropy distributions" (http://www.harremoes.dk/Peter/poisson.pdf) (PDF). *IEEE Transactions on Information Theory*. **47** (5): 2039–2041. doi:10.1109/18.930936 (https://doi.org/10.1109%2F18.930936).
7. Shepp, Lawrence; Olkin, Ingram (1981). "Entropy of the sum of independent Bernoulli random variables and of the multinomial distribution" (https://books.google.co.uk/books?id=9qziBQAAQBAJ). In Gani, J.; Rohatgi, V.K. (eds.). *Contributions to probability: A collection of papers dedicated to Eugene Lukacs*. New York: Academic Press. pp. 201–206. ISBN 0-12-274460-8. MR 0618689 (https://www.ams.org/mathscinet-getitem?mr=0618689).
8. Hillion, Erwan; Johnson, Oliver (2015-03-05). "A proof of the Shepp-Olkin entropy concavity conjecture". *Bernoulli*. **23**: 3638–3649. arXiv:1503.01570 (https://arxiv.org/abs/1503.01570). doi:10.3150/16-BEJ860 (https://doi.org/10.3150%2F16-BEJ860).