I'm not a bot

Share — copy and redistribute the material in any medium or format for any purpose, even commercially. Adapt — remix, transform, and build upon the material for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms. Attribution — You must give appropriate credit , provide a link to the license, and indicate if changes were made . You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original. No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits. You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation . No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material. Ejercicios 1 2 3 4 5 6 7 8 9 10 11 12 Tenemos una distribución binomial $B(50; 0{,}15)$, siendo X la variable que expresa el número de personas que han leído más de tres libros. Como el número de datos es muy grande y se verifica que $n \cdot p = 7{,}5 \geq 5$ y $n \cdot q = 42{,}5 \geq 5$, aproximaremos mediante una distribución normal de parámetros : a) Sea X la variable aleatoria que describe el número de veces que sal el número 4. Se trata de una $B(6; 0{,}1)$ y la probabilidad pedida es la siguiente : b) Consideramos ahora que la variable X describe el número de veces que sale el número 5. Se trata así de una distribución $N(150; 0{,}1)$. Como $n \cdot p = 15 \geq 5$ y $n \cdot q = 135 \geq 5$ podemos aproximar nuestra distribución a una normal de parámetros : La variable X = " número de viviendas con internet " sigue una distribución binomial $B(80; 0{,}25)$. Como $n \cdot p = 80 \cdot 0{,}25 = 20 > 5$ y $n \cdot q = 80 \cdot 0{,}75 = 60 > 5$, se puede aproximar X por una normal $N(20; 3{,}87)$. X = " número de pacientes que deja de fumar " $p = 0{,}8 \rightarrow q = 1 - p = 0{,}2$ n = " número de pacientes que se ha sometido a dicho régimen " = 100 Tenemos una distribución binomial $B(100; 0{,}8)$. Consideramos la variable X, que indica el número de caras que salen en los 120 lanzamientos de la moneda. X es, por tanto, una distribución binomial de parámetros $n = 120$ y $p = 1/3$. Como n es grande, $n \cdot p = 40 > 5$ y $n \cdot q = 80 > 5$, X se puede aproximar a una distribución normal de parámetros siguientes: La distribución binomial $B(n, p)$ se puede aproximar a una distribución normal $N(\mu, \sigma)$ en el caso donde n es grande y ni p ni q son próximos a cero. La aproximación será por tanto mejor cuanto mayor sea n. Se puede considerar que la aproximación es buena si $n > 30$ y p toma valores próximos a 0,5. También se suele considerar que la aproximación es buena si $n \cdot p$ y $n \cdot q$ son iguales o superiores a 5. Calculator-online.net is your ultimate destination for a wide range of free online calculators. These tools include AI tools, health, finance, statistics, maths, physics, and chemistry. Everyone deserves instant and free access to reliable calculations. Our mission is to provide accurate and up-to-date results to tackle challenges with precision. Why does Calculator Online strive to be a one-stop destination? The team of online calculators continues to update, and new tools are being added regularly. Whether you are tackling basic maths problems or delving into more complex equations, this platform is designed to streamline your mathematical endeavours. Every tool developed by the Calculator Online is designed for accurate results, showing our strong commitment to your satisfaction. If you feel any inconvenience regarding these online calculators, feel free to Contact Us. Increase your mathematical productivity with the Calculator Online, your best companion in mathematics. An online Binomial Distribution Calculator can find the cumulative and binomial probabilities for the given values. Now, you can determine the standard deviation, variance, and mean of the binomial distribution quickly with a binomial probability distribution calculator. In the following article, you can understand what exactly is the binomial distribution, when and how to apply it, and much more information that you should know about the probability distribution. Let’s begin with some basics! What is Binomial Distribution? In statistics, the binomial distribution is a discrete probability distribution that only gives two possible results in an experiment either failure or success. For example, if we toss with a coin, there can only be two possible outcomes: tails or heads, and when taking any test, there can only be two outcomes: pass or fail. This distribution is called the binomial probability distribution. Two parameters p and n are used in the binomial distribution. The variable "n" represents the frequency of the experiment, and the variable "p" represents the probability of the result. Assuming that the dice is randomly rolled 10 times, then the probability of each roll is 2. If you roll the dice 10 times, you will get a binomial distribution with p = ⅙ and n = 10. Explore the formula for calculating the distribution of two results in multiple experiments. Binomial Distribution Formula: The formula for the binomial distribution is: $$ P(x) = pr (1 − p) n−r . nCr $$ Or, $$ P(x) = pr (1 − p) n−r . [n!/r!(n−r)!] $$ Where, r = Total number of successful trails n = Total number of events p = Probability of success 1 – p = Probability of failure nCr = [n!/r!(n−r)]! However, an online Poisson Distribution Calculator determines the probability of the event happening many times over some given intervals. How to Calculate Binomial Probability Distribution? Here’s a comprehensive example that describes how a binomial distribution calculator works which can be helpful for determining the binomial distribution manually if required. Example: A coin is tossed 5 times with 0.13 probability for the number of successes (x) and the condition with exactly X success P(X = x). Solution: Probability of exactly 3 successes $$P(X = 3) = 0.016629093$$ Use a binomial CDF calculator to get the standard deviation, variance, and mean of binomial distribution based on the number of trails you provided. Mean: μ = np = ((5) × (0.13)) = 0.65 Variance: σ2 = np (1 − p) = (5) (0.13) (1 − 0.13) = 0.5655 Standard deviation: σ = np(1 − p) = (5) (0.13) (1 − 0.13) = 0.75199734042083 Given Values : Trials =5, p = 0.13 and X = 3 Formula: $$ P(X) = (nX) · pX · (1 − p)^{n – X} $$ The binomial coefficient, (nX) is defined by: $$ (nX) = n! / X! (n−X)! $$ The binomial probability formula that is used by the binomial probability calculator with the binomial coefficient is: $$ P(X) = n! / X! (n − X)! · p^X· (1 − p) n − X $$ Where, n = number of trials p = probability of success on a single trial, X = number of successes Substituting in values for this problem, n = 5, p = 0.13 and X = 3: $$ P (3) = 5! / 3! (5−3)! · 0.133 · (1 − 0.13) 5 − 3 $$ After Solving the expression: $$ P (3) = 0.016629093 $$ The Binomial Distribution Calculator Provide a table for: n = 5, p = 0.13 $$ P(0) = 0.4984209207 $$ $$ P(1) = 0.3723834465 $$ $$ P(2) = 0.111287007 $$ $$ P(3) = 0.016629093 $$ $$ P(4) = 0.0012424035 $$ $$ P(5) = 3.71293E−5 $$ Pie Chart for Probability Relative: The binomial probability calculator displays a pie chart for probability relative: Probability vs Number of successes Graph: However, an online Binomial Theorem Calculator helps you to find the expanding binomials for the given binomial equation. Binomial Probability Distribution: In the probability distribution, the number of "successes" in the sequence of n experiments, where every time is asking for "yes or no", then the result is expressed as a Boolean value for success/Yes/True/probability p or failure/no/false/probability q = 1-p. The successful/failed unit test is also called the Bernoulli test or Bernoulli experiment and the series of results is called the Bernoulli process. For n = 1 that is for a single experiment, the binomial distribution is the Bernoulli distribution. The binomial distribution is the basis of the famous binomial statistical significance test. Negative Binomial Distribution: In probability, the number of successful results in a series of identically distributed and independent distributed Bernoulli tests before a certain number of failures occur. This is called a negative binomial distribution. The number of failures/errors is represented by the letter "r". Binomial Distribution Mean and Variance: For the binomial distribution, the variance, mean, and standard deviation of a given number of successes are expressed by the following formula $$ Variance, σ2 = npq $$ $$ Mean, μ = np $$ $$ Standard Deviation σ= √(npq) $$ These formulae are used by a binomial distribution calculator for determining the variance, mean, and standard deviation. Where, p = probability of success q = probability of failure Binomial and Normal Distribution: The main difference between the normal distribution and the binomial distribution is that the binomial distribution is discrete, while the normal distribution is continuous. It means the binomial distribution is the limited number of events whereas the normal distribution has an infinite number of events. If the sample size of the binomial distribution is very large, then the distribution curve of the binomial distribution is the same as the normal distribution curve. Properties of Binomial Distribution: The main properties of the binomial distribution are: There are two possible outcomes: success or failure, true or false, yes or no. There is 'n' fixed number of n repeated attempts or the independent trails. Each trail has a different probability of failure or success. Only count the number of successes n that are independent trials. Each test is an independent trial that means the result of one trial has no effect on the result of another trial. How does Binomial Distribution Calculator Works? An online binomial probability distribution calculator finds the probabilities for different conditions by using these steps: Input: First, enter the number of trails, probability, and the number of successes. Now, choose the condition for determining the binomial distribution. Click on the Calculate button. Output: The binomial probability formula calculator displays the variance, mean, and standard deviation. An online binomial calculator shows the binomial coefficients, binomial distribution table, pie chart, and bar graph for probability and number of success. FAQ: What are the criteria of binomial distribution? The number of trials/tests should be specified. Each trial must be independent. The probability of success in each trial is the same. Uses of Binomial Distribution in real life. In real life, you can find many examples of binomial distributions. For example, when a new medicine is used to treat a disease, it either cures the disease (which is successful) or cannot cure the disease (which is a failure). Conclusion: Use this online binomial distribution calculator to evaluate the cumulative probabilities for the binomial distribution, given the number of trials (n), the number of success (X), and the probability (p) of the successful outcomes occurring. It also computes the variance, mean of binomial distribution, and standard deviation with different graphs. Reference: From the source of Wikipedia: Probability mass function, Cumulative distribution function, Expected value and variance, Higher moments, Sums of binomials, Ratio of two binomial distributions. From the source of Investopedia: Analyzing Binomial Distribution, probability distribution, normal distribution, binomial distribution. From the source of Lumen Learning: Binomial Probability Distribution, Concept Review, Formula Review. Probability distribution "Binomial model" redirects here. For the binomial model in options pricing, see Binomial options pricing model. Binomial distribution Probability mass function Cumulative distribution functionNotation $B(n, p)$ Parameters $n \in \{0, 1, 2, \ldots\}$ – number of trials $p \in [0, 1]$ – success probability for each trial $q = 1 - p$ $q=1-p$ Support $k \in \{0, 1, \ldots, n\}$ – number of successesPMF $\binom{n}{k} p^k q^{n-k}$ CDF $I_q(n - \lfloor k \rfloor, 1 + \lfloor k \rfloor)$ (regularized incomplete beta function)Mean np Median $\lfloor np \rfloor$ or $\lceil np \rceil$ Mode $\lfloor (n+1)p \rfloor$ or $\lceil np \rceil - 1$ Variance $np(1-p)$ Skewness $\frac{q-p}{\sqrt{npq}}$ Excess kurtosis $\frac{1-6pq}{npq}$ Entropy $\frac{1}{2}\log_2(2\pi enpq)+O(\frac{1}{n})$ in shannons. For nats, use the natural log in the log.MGF $(q + pe^t)^n$ $(q+pe^{t})^{n}$ CF $(q + pe^{it})^n$ $(q+pe^{it})^{n}$ PGF $G(z) = [q + pz]^n$ $G(z)=[q+pz]^{n}$ Fisher information $g_n(p) = n/p(1-p)$ $g_{n}(p)=\frac{n}{pq}$ (for fixed n) Part of a series on statisticsProbability theory Probability Axioms Determinism System Indeterminism Randomness Probability space Sample space Event Collectively exhaustive events Elementary event Mutual exclusivity Outcome Singleton Experiment Bernoulli trial Probability distribution Bernoulli distribution Binomial distribution Exponential distribution Normal distribution Pareto distribution Poisson distribution Probability measure Random variable Bernoulli process Continuous or discrete Expected value Variance Markov chain Observed value Random walk Stochastic process Complementary event Joint probability Marginal probability Conditional probability Independence Conditional independence Law of total probability Law of large numbers Bayes' theorem Boole's inequality Venn diagram Tree diagram vte Binomial distribution for $p = 0.5$with n and k as in Pascal's triangleThe probability that a ball in a Galton box with 8 layers ($n = 8$) ends up in the central bin ($k = 4$) is 70/256. In probability theory and statistics, the binomial distribution with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent experiments, each asking a yes–no question, and each with its own Boolean-valued outcome: success (with probability p) or failure (with probability $q = 1 - p$). A single success/failure experiment is also called a Bernoulli trial or Bernoulli experiment, and a sequence of outcomes is called a Bernoulli process; for a single trial, i.e., $n = 1$, the binomial distribution is a Bernoulli distribution. The binomial distribution is the basis for the binomial test of statistical significance.[1] The binomial distribution is frequently used to model the number of successes in a sample of size n drawn with replacement from a population of size N. If the sampling is carried out without replacement, the draws are not independent and so the resulting distribution is a hypergeometric distribution, not a binomial one. However, for N much larger than n, the binomial distribution remains a good approximation, and is widely used. If the random variable X follows the binomial distribution with parameters $n \in \mathbb{N}$ and $p \in [0, 1]$, we write $X \sim B(n, p)$. The probability of getting exactly k successes in n independent Bernoulli trials (with the same rate p) is given by the probability mass function: $f(k, n, p) = \Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k = 0, 1, 2, \ldots$, where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ is the binomial coefficient. The formula can be understood as follows: $pk qn−k$ is the probability of obtaining the sequence of n independent Bernoulli trials in which k trials are "successes" and the remaining $n - k$ trials result in "failure". Since the trials are independent with probabilities remaining constant between them, any sequence of n trials with k successes (and $n - k$ failures) has the same probability of being achieved (regardless of positions of successes within the sequence). There are $\binom{n}{k}$ such sequences, since the binomial coefficient $\binom{n}{k}$ counts the number of ways to choose the positions of the k successes among the n trials. The binomial distribution is concerned with the probability of obtaining any of these sequences, meaning the probability of obtaining one of them ($pk qn−k$) must be added $\binom{n}{k}$ times, hence $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. In creating reference tables for binomial distribution probability, usually, the table is filled in up to n/2 values. This is because for $k > n/2$, the probability can be calculated by its complement as $f(k, n, p) = f(n - k, n, 1 - p)$. {\displaystyle f(k,n,p)=f(n-k,n,1-p).} Looking at the expression f(k, n, p) as a function of k, there is a k value that maximizes it. This k value can be found by calculating $\frac{f(k+1,n,p)}{f(k,n,p)}$ and comparing it to 1. There is always an integer M that satisfies $(n + 1) p - 1 \leq M < (n + 1) p$. {\displaystyle (n+1)p-1\leq M M, with the exception of the case where $(n + 1)p$ is an integer. In this case, there are two values for which f is maximal: $(n + 1) p$ and $(n + 1) p - 1$. M is the most probable outcome (that is, the most likely, although this can still be unlikely overall) of the Bernoulli trials and is called the mode. Equivalently, $M - p < np \leq M + 1 - p$. Taking the floor function, we obtain $M = \operatorname{floor}(np)$.[note 1] Suppose a biased coin comes up heads with probability 0.3 when tossed. The probability of seeing exactly 4 heads in 6 tosses is $f(4, 6, 0.3) = \binom{6}{4} 0.3^4 (1-0.3)^{6-4} = 0.059535$. {\displaystyle f(4,6,0.3)=\binom{6}{4}0.3^{4}(1-0.3)^{6-4}=0.059535.} The cumulative distribution function can be expressed as: $F(k; n, p) = \Pr(X \leq k) = \sum_{i=0}^{\lfloor k \rfloor} \binom{n}{i} p^i (1-p)^{n-i}$, {\displaystyle F(k;n,p)=\Pr(X\leq k)=\sum _{i=0}^{\lfloor k\rfloor }\choose i}p^{i}(1-p)^{n-i},} where $\lfloor k \rfloor$ is the "floor" under k, i.e. the greatest integer less than or equal to k. It can also be represented in terms of the regularized incomplete beta function, as follows:[3] $F(k; n, p) = \Pr(X \leq k) = I_{1-p}(n-k, k+1) = (n - k) \binom{n}{k} \int_0^{1-p} t^{n-k-1} (1-t)^k \, dt$, {\displaystyle {\begin{aligned}F(k;n,p)&=\Pr(X\leq k)\\&=I_{1-p}(n-k,k+1)\\&=(n-k){n \choose k}\int _{0}^{1-p}t^{n-k-1}(1-t)^{k}\,dt,\end{aligned}}} which is equivalent to the cumulative distribution functions of the beta distribution and of the F-distribution:[4] $F(k; n, p) = F_{\text{beta-distribution}}(x = 1 - p; \alpha = n - k, \beta = k + 1)$ {\displaystyle F(k;n,p)=F_{\text{beta-distribution}}\left(x=1-p;\alpha =n-k,\beta =k+1\right)} $F(k; n, p) = F_{F\text{-distribution}}(x = \frac{1-p}{p} \frac{k+1}{n-k}; d_1 = 2(n-k), d_2 = 2(k+1))$. {\displaystyle F(k;n,p)=F_{F{\text{-distribution}}}\left(x={\frac {1-p}{p}}{\frac {k+1}{n-k}};d_{1}=2(n-k),d_{2}=2(k+1)\right).} Some closed-form bounds for the cumulative distribution function are given below. If $X \sim B(n, p)$, that is, X is a binomially distributed random variable, n being the total number of experiments and p the probability of each experiment yielding a successful result, then the expected value of X is:[5] $E[X] = np$. {\displaystyle \operatorname {E}[X]=np.} This follows from the linearity of the expected value along with the fact that X is the sum of n identical Bernoulli random variables, each with expected value p. In other words, if $X_1, \ldots, X_n$ {\displaystyle X_{1},\ldots ,X_{n}} are identical (and independent) Bernoulli random variables with parameter p, then $X = X_1 + \cdots + X_n$ $E[X] = E[X_1 + \cdots + X_n] = p + \cdots + p = np$. The variance is: $\operatorname{Var}(X) = npq = np(1-p)$. {\displaystyle \operatorname {Var}(X)=npq=np(1-p).} This similarly follows from the fact that the variance of a sum of independent random variables is the sum of the variances. The first 6 central moments, defined as $\mu_c = E[(X - E[X])^c]$ {\displaystyle \mu _{c}=\operatorname {E}\left[(X-\operatorname {E}[X])^{c}\right]} , are given by $\mu_1 = 0$, $\mu_2 = np(1-p)$, $\mu_3 = np(1-p)(1-2p)$, $\mu_4 = np(1-p)(1 + (3 n - 6) p(1-p))$, $\mu_5 = np(1-p)(1-2p)(10 - 12 p(1-p))$, $\mu_6 = np(1-p)(1-30p(1-p)(1-4p(1-p)) + 5np(1-p)(5-26p(1-p)) + 15n^2p^2(1-p)^2$. {\displaystyle {\begin{aligned}\mu _{1}&=0,\\\mu _{2}&=np(1-p),\\\mu _{3}&=np(1-p)(1-2p),\\\mu _{4}&=np(1-p)(1+(3n-6)p(1-p)),\\\mu _{5}&=np(1-p)(1-2p)(10-12)p(1-p)),\\\mu _{6}&=np(1-p)(1-30p(1-p)(1-4p(1-p))+5np(1-p)(5-26p(1-p))+15n^{2}p^{2}(1-p)^{2}).\end{aligned}}} The non-central moments satisfy $E[X] = np(1-p) + n^2 p^2$. {\displaystyle {\begin{aligned}\operatorname {E}[X]&=np(1-p)+n^{2}p^{2},\end{aligned}}} and in general[6][7] $E[X^c] = \sum_{k=0}^{c} \left\{ {c \atop k} \right\} n^{\underline{k}} p^k$, where $\left\{ {c \atop k} \right\}$ are the Stirling numbers of the second kind, and $n^{\underline{k}} = n(n-1)\cdots(n-k+1)$ is the k {\displaystyle k} th falling power of n {\displaystyle n}. A simple bound [8] follows by bounding the Binomial moments via the higher Poisson moments: $E[X^c] \leq \left(\frac{c}{\ln(c/(np)+1)}\right)^c \exp(np)$. {\displaystyle \operatorname {E}[X^{c}]\leq \left({\frac {c}{\ln(c/(np)+1)}}\right)^{c}\leq (np)^{c}\exp \left({\frac {c^{2}}{2np}}\right).} This shows that if $c = O(np)$ {\displaystyle c=O({\sqrt {np}})} , then $E[X^c]$ {\displaystyle \operatorname {E}[X^{c}]} is at most a constant factor away from $E[X]^c$ {\displaystyle \operatorname {E}[X]^{c}} Usually the mode of a binomial B(n, p) distribution is equal to $\lfloor (n + 1) p \rfloor$ {\displaystyle \lfloor (n+1)p\rfloor } , where $\lfloor \cdot \rfloor$ is the floor function. However, when $(n + 1)p$ is an integer and p is neither 0 nor 1, then the distribution has two modes: $(n + 1)p$ and $(n + 1)p - 1$. When p is equal to 0 or 1, the mode will be 0 and n correspondingly. These cases can be summarized as follows: mode = $\lfloor (n + 1) p \rfloor$ if $(n + 1) p$ is 0 or a noninteger, $(n + 1) p$ and $(n + 1) p - 1$ if $(n + 1) p \in \{1, \ldots, n\}$, n if $(n + 1) p = n + 1$. {\displaystyle {\text{mode}}=\begin{cases}\lfloor (n+1)\,p\rfloor &{\text{if }}(n+1)p\neq 1,\\ (n+1)\,p\ {\text{ and }}\ (n+1)\,p-1&{\text{if }}(n+1)p=1\\ n&{\text{if }}(n+1)p=n+1.\end{cases}}} Proof: Let $f(k) = \binom{n}{k} p^k q^{n-k}$. {\displaystyle f(k)={\binom {n}{k}}p^{k}q^{n-k}.} For $p = 0$ {\displaystyle p=0} only $f(0)$ {\displaystyle f(0)} has a nonzero value with $f(0) = 1$ {\displaystyle f(0)=1} . For $p = 1$ {\displaystyle p=1} we find $f(n) = 1$ {\displaystyle f(n)=1} and $f(k) = 0$ {\displaystyle f(k)=0} for $k \neq n$ {\displaystyle k\neq n} . This proves that the mode is 0 for $p = 0$ {\displaystyle p=0} and n {\displaystyle n} for $p = 1$ {\displaystyle p=1} . Let $0 < p < 1$ {\displaystyle 0 f(k) {\displaystyle {\begin{aligned}k>(n+1)p-1\Rightarrow f(k+1) 10 and n1 ≠ 0, n.[23] See here for $n \leq 10$ {\displaystyle n\leq 10} .[24] For n1 = 0, n use the Wilson (score) method below. Main article: Binomial proportion confidence interval § Arcsine transformation [25] sin 2 (arcsin (p ^) ± z 2 n) . {\displaystyle \sin ^{2}\left(\arcsin \left({\sqrt {\widehat {p\,}}}\right)\pm {\frac {z}{2{\sqrt {n}}}}\right).} Main article: Binomial proportion confidence interval § Wilson score interval The notation in the formula below differs from the previous formulas in two respects:[26] Firstly, zx has a slightly different interpretation in the formula below: it has its ordinary meaning of 'the xth quantile of the standard normal distribution', rather than being a shorthand for 'the (1 − x)th quantile'. Secondly, this formula does not use a plus-minus to define the two bounds. Instead, one may use $z = z_{\alpha /2}$ {\displaystyle z=z_{\alpha /2}} to get the lower bound, or use $z = z_{1-\alpha /2}$ {\displaystyle z=z_{1-\alpha /2}} to get the upper bound. For example: for a 95% confidence level the error α {\displaystyle \alpha } = 0.05, so one gets the lower bound by using $z = z_{\alpha /2} = z_{0.025} = -1.96$ {\displaystyle z=z_{\alpha /2}=z_{0.025}=-1.96} , and one gets the upper bound by using $z = z_{1-\alpha /2} = z_{0.975} = 1.96$ {\displaystyle z=z_{1-\alpha /2}=z_{0.975}=1.96} . $p^+ \pm z \sqrt{\frac{\widehat{p\,}^+(1-\widehat{p\,}^+)}{n}} + \frac{z^2}{2n} = \frac{\widehat{p\,}^+ + \frac{z^2}{2n} \pm z \sqrt{\frac{\widehat{p\,}^+(1-\widehat{p\,}^+)}{n} + \frac{z^2}{4n^2}}}{1+\frac{z^2}{n}}$. [27] The so-called "exact" (Clopper–Pearson) method is the most conservative.[21] (Exact does not mean perfectly accurate; rather, it indicates that the estimates will not be less conservative than the true value.) The Wald method, although commonly recommended in textbooks, is the most biased. [clarification needed] If $X \sim B(n, p)$ and $Y \sim B(m, p)$ are independent binomial variables with the same probability p, then $X + Y$ is again a binomial variable; its distribution is $Z = X + Y \sim B(n + m, p)$:[28] $P(Z = k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i} = \binom{n+m}{k} p^k (1-p)^{n+m-k}$ {\displaystyle {\begin{aligned}\operatorname {P}(Z=k)&=\sum _{i=0}^{k}\left[{\binom {n}{i}}p^{i}(1-p)^{n-i}\right]\left[{\binom {m}{k-i}}p^{k-i}(1-p)^{m-k+i}\right]\\&={\binom {n+m}{k}}p^{k}(1-p)^{n+m-k}\end{aligned}}} A Binomial distributed random variable $X \sim B(n, p)$ can be considered as the sum of n Bernoulli distributed random variables. So the sum of two Binomial distributed random variables $X \sim B(n, p)$ and $Y \sim B(m, p)$ is equivalent to the sum of $n + m$ Bernoulli distributed random variables, which means $Z = X + Y \sim B(n + m, p)$. This can also be proven directly using the addition rule. However, if X and Y do not have the same probability p, then the variance of the sum will be smaller than the variance of a binomial variable distributed as $B(n + m, p)$. The binomial distribution is a special case of the Poisson binomial distribution, which is the distribution of a sum of n independent non-identical Bernoulli trials B(pi).[29] This result was first derived by Katz and coauthors in 1978.[30] Let $X \sim B(n, p1)$ and $Y \sim B(m, p2)$ be independent. Let $T = (X/n) / (Y/m)$. Then $\log(T)$ is approximately normally distributed with mean $\log(p1/p2)$ and variance $((1/p1) - 1)/n + ((1/p2) - 1)/m$. If $X \sim B(n, p)$ and $Y | X \sim B(X, q)$ (the conditional distribution of Y, given X), then Y is a simple binomial random variable with distribution $Y \sim B(n, pq)$. For example, imagine throwing n balls to a basket UX and taking the balls that hit and throwing them to another basket UY. If p is the probability to hit UX then $X \sim B(n, p)$ is the number of balls that hit UX. If q is the probability to hit UY then the number of balls that hit UY is $Y \sim B(X, q)$ and therefore $Y \sim B(n, pq)$. [Proof] Since $X \sim B(n, p)$ {\displaystyle X\sim B(n,p)} and $Y \sim B(X, q)$ {\displaystyle Y\sim B(X,q)} , by the law of total probability, $\Pr[Y = m] = \sum_{k=m}^{n} \Pr[Y = m \mid X = k] \Pr[X = k] = \sum_{k=m}^{n} \binom{n}{k} \binom{k}{m} p^k q^m (1-p)^{n-k} (1-q)^{k-m}$ {\displaystyle {\begin{aligned}\Pr[Y=m]&=\sum _{k=m}^{n}\Pr[Y=m\mid X=k]\Pr[X=k]\\[2pt]&=\sum _{k=m}^{n}{\binom {n}{k}}{\binom {k}{m}}p^{k}q^{m}(1-p)^{n-k}(1-q)^{k-m}\end{aligned}}} Since $\binom{n}{k} \binom{m}{k} = \binom{n}{m} \binom{n-m}{k-m}$, {\displaystyle {\binom {n}{k}}{\binom {k}{m}}={\binom {n}{m}}{\binom {n-m}{k-m}},} the equation above can be expressed as $\Pr[Y = m] = \sum_{k=m}^{n} \binom{n}{m} \binom{n-m}{k-m} p^k q^m (1-p)^{n-k} (1-q)^{k-m}$ {\displaystyle \Pr[Y=m]=\sum _{k=m}^{n}{\binom {n}{m}}{\binom {n-m}{k-m}}p^{k}q^{m}(1-p)^{n-k}(1-q)^{k-m}} Factoring $p^k = p^m p^{k-m}$ {\displaystyle p^{k}=p^{m}p^{k-m}} and pulling all the terms that don't depend on k {\displaystyle k} out of the sum now yields $\Pr[Y = m] = \binom{n}{m} p^m q^m (\sum_{k=m}^{n} \binom{n-m}{k-m} p^{k-m} (1-p)^{n-k} (1-q)^{k-m})$ = (n m) (p q) m (�sum_{k=m}^{n} \binom{n-m}{k-m} p^{k-m} (1-p)^{n-k} (1-q)^{k-m} \right) = (n m) (p q) m (∑ k = m − n (n − m k − m) (p (1 − q)) k − m (1 − p) n − k) {\displaystyle {\begin{aligned}\Pr[Y=m]&=\left(\sum _{k=m}^{n}{\binom {n-m}{k-m}}p^{k-m}(1-p)^{n-k}(1-q)^{k-m}\right)\\[2pt]&=\left(\sum _{k=m}^{n}{\binom {n-m}{k-m}}(pq)^{m}\right)(1-p)^{n-k}\right)\end{aligned}}} After substituting $i = k - m$ {\displaystyle i=k-m} in the expression above, we get $\Pr[Y = m] = \binom{n}{m} (pq)^m \sum_{i=0}^{n-m} \binom{n-m}{i} (p-pq)^i (1-p)^{n-m-i}$ {\displaystyle \Pr[Y=m]=\left(\sum _{i=0}^{n-m}{\binom {n-m}{i}}(p-pq)^{i}(1-p)^{n-m-i}\right)} Notice that the sum (in the parentheses) above equals $(p - pq + 1 - p)^{n-m}$ {\displaystyle (p-pq+1-p)^{n-m}} by the binomial theorem. Substituting this in finally yields $\Pr[Y = m] = \binom{n}{m} (pq)^m (p - pq + 1 - p)^{n-m} = \binom{n}{m} (pq)^m (1-p)^n$ {\displaystyle {\begin{aligned}\Pr[Y=m]&={\binom {n}{m}}(pq)^{m}(p-pq+1-p)^{n-m}\\[4pt]&={\binom {n}{m}}(pq)^{m}(1-p)^{n-m}\end{aligned}}} and thus $Y \sim B(n, pq)$ {\displaystyle Y\sim B(n,pq)} as desired. The Bernoulli distribution is a special case of the binomial distribution, where $n = 1$. Symbolically, $X \sim B(1, p)$ has the same meaning as $X \sim \operatorname{Bernoulli}(p)$. Conversely, any binomial distribution, $B(n, p)$, is the distribution of the sum of n independent Bernoulli trials, $\operatorname{Bernoulli}(p)$, each with the same probability p.[31] See also: Binomial proportion confidence interval § Normal approximation interval Binomial probability mass function and normal probability density function approximation for $n = 6$ and $p = 0.5$ If n is large enough, then the skew of the distribution is not too great. In this case a reasonable approximation to $B(n, p)$ is given by the normal distribution $N(np, np(1-p))$, {\displaystyle {\mathcal {N}}(np,\,np(1-p)),} and this basic approximation can be improved in a simple way by using a suitable continuity correction. The basic approximation generally improves as n increases (at least 20) and is better when p is not near to 0 or 1.[32] Various rules of thumb may be used to decide whether n is large enough, and p is far enough from the extremes of zero or one: One rule[32] is that for $n > 5$ the normal approximation is adequate if the absolute value of the skewness is strictly less than 0.3; that is, if $\left|\frac{1-2p}{\sqrt{np(1-p)}}\right| < 0.3$. {\displaystyle {\frac {|1-2p|}{\sqrt {np(1-p)}}}=\left|{\frac {1}{\sqrt {n}}}\left({\frac {1-p}{p}}-{\frac {p}{1-p}}\right)\right| 9 (1 − p p) and n > 9 (p 1 − p) . {\displaystyle n>9\left({\frac {1-p}{p}}\right)\quad {\text{and}}\quad n>9\left({\frac {p}{1-p}}\right).} [Proof] The rule $np \pm 3\sqrt{np(1-p)} \in (0, n)$ {\displaystyle np\pm 3{\sqrt {np(1-p)}}\in (0,n)} is totally equivalent to request that $np - 3\sqrt{np(1-p)} > 0$ and $np + 3\sqrt{np(1-p)} < n$. {\displaystyle np-3{\sqrt {np(1-p)}}>0\quad {\text{and}}\quad np+3{\sqrt {np(1-p)}} 3 n p (1 − p) and n (1 − p) > 3 n p (1 − p) . {\displaystyle np>3{\sqrt {np(1-p)}}\quad {\text{and}}\quad n(1-p)>3{\sqrt {np(1-p)}}.} Since $0 < p < 1$ {\displaystyle 0 9 (p 1 − p) . {\displaystyle n>9\left({\frac {1-p}{p}}\right)\quad {\text{and}}\quad n>9\left({\frac {p}{1-p}}\right).} Notice that these conditions automatically imply that $n > 9$ {\displaystyle n>9} . On the other hand, apply again the square root and divide by 3, $n 3 > 1 - p p$ $p > 0$ and $n 3 > p 1 - p > 0$. {\displaystyle {\frac {\sqrt {n}}{3}}>{\sqrt {\frac {1-p}{p}}}>0\quad {\text{and}}\quad {\frac {\sqrt {n}}{3}}>{\sqrt {\frac {p}{1-p}}}>0.} Subtracting the second set of inequalities from the first one yields: $n 3 > 1 - p p - p 1 - p > - n 3$; {\displaystyle {\frac {\sqrt {n}}{3}}>{\sqrt {\frac {1-p}{p}}}-{\sqrt {\frac {p}{1-p}}}>-{\frac {\sqrt {n}}{3}};} and so, the desired first rule is satisfied, $\left|1 - p p - p 1 - p\right| < n 3$. {\displaystyle \left|{\sqrt {\frac {1-p}{p}}}-{\sqrt {\frac {p}{1-p}}}\,\right|9(1-p)\quad {\text{and}}\quad n(1-p)\geq 9>9p.} We only have to divide now by the respective factors p {\displaystyle p} and $1 - p$ {\displaystyle 1-p} , to deduce the alternative form of the 3-standard-deviation rule: $n > 9 (1 − p p)$ and $n > 9 (p 1 − p)$. {\displaystyle n>9\left({\frac {1-p}{p}}\right)\quad {\text{and}}\quad n>9\left({\frac {p}{1-p}}\right).} The following is an example of applying a continuity correction. Suppose one wishes to calculate $\Pr(X \leq 8)$ for a binomial random variable X. If Y has a distribution given by the normal approximation, then $\Pr(X \leq 8)$ is approximated by $\Pr(Y \leq 8.5)$. The addition of 0.5 is the continuity correction; the uncorrected normal approximation gives considerably less accurate results. This approximation, known as de Moivre–Laplace theorem, is a huge time-saver when undertaking calculations by hand (exact calculations with large n are very onerous); historically, it was the first use of the normal distribution, introduced in Abraham de Moivre's book The Doctrine of Chances in 1738. Nowadays, it can be seen as a consequence of the central limit theorem since B(n, p) is a sum of n independent, identically distributed Bernoulli variables with parameter p. This fact is the basis of a hypothesis test, a "proportion z-test", for the value of p using x/n, the sample proportion and estimator of p, in a common test statistic.[35] For example, suppose one randomly samples n people out of a large population and ask them whether they agree with a certain statement. The proportion of people who agree will of course depend on the sample. If groups of n people were sampled repeatedly and truly randomly, the proportions would follow an approximate normal distribution with mean equal to the true proportion p of agreement in the population and with standard deviation $\sigma = \sqrt{\frac{p(1-p)}{n}}$ The binomial distribution converges towards the Poisson distribution as the number of trials goes to infinity while the product np converges to a finite limit. Therefore, the Poisson distribution with parameter $\lambda = np$ can be used as an approximation to B(n, p) of the binomial distribution if n is sufficiently large and p is sufficiently small. According to rules of thumb, this approximation is good if $n \geq 20$ and $p \leq 0.05$[36] such that $np \leq 1$, or if $n > 50$ and $p < 0.1$ such that $np < 5$,[37] or if $n \geq 100$ and $np \leq 10$.[38][39] Concerning the accuracy of Poisson approximation, see Novak,[40] ch. 4, and references therein. Poisson limit theorem: As n approaches ∞ and p approaches 0 with the product np held fixed, the Binomial(n, p) distribution approaches the Poisson distribution with expected value $\lambda = np$.[38] de Moivre–Laplace theorem: As n approaches ∞ while p remains fixed, the distribution of $X - np \sqrt{np(1-p)}$ {\displaystyle {\frac {X-np}{\sqrt {np(1-p)}}}} approaches the normal distribution with expected value 0 and variance 1. This result is sometimes loosely stated by saying that the distribution of X is asymptotically normal with expected value 0 and variance 1. This result is a specific case of the central limit theorem. The binomial distribution and beta distribution are different views of the same model of repeated Bernoulli trials. The binomial distribution is the PMF of k successes given n independent events each with a probability p of success. Mathematically, when $\alpha = k + 1$ and $\beta = n - k + 1$, the beta distribution and the binomial distribution are related by[clarification needed] a factor of $n + 1$: $\operatorname{Beta}(p; \alpha, \beta) = (n + 1) B(k, n; p)$ {\displaystyle \