



Addition of two Binomial Distribution

What is the distribution of the variable X given

$$X = Y + Z$$

, where $Y \sim \text{Binomial}(n, P_Y)$ and $Z \sim \text{Binomial}(n, P_Z)$?

For the special case, when $P_Y = P_Z = P$, I think that $X \sim \text{Binomial}(2n, P)$ is correct. If $P_A \neq P_B$, the distribution might eventually just be $\text{Binomial}\left(2n, \frac{P_A + P_B}{2}\right)$ but I can't prove it.

If the problem is more complicated than I expect and we can't derive the whole distribution, can we tell something about the mean and the variance of X ?

(probability) (probability-theory) (probability-distributions) (random-variables)

asked Feb 17 '15 at 20:41



[Remi.b](#)

581 6 13

Assuming independence, means and variance should be additive. – [jameselmore](#) Feb 17 '15 at 20:44

2 @jameselmore Additivity of the means is unrelated to independence. – [Did](#) Feb 17 '15 at 22:18

4 Answers

Assuming Y and Z are independent, $X = Y + Z$ has mean $E[Y] + E[Z] = nP_Y + nP_Z$ and variance $\text{Var}(Y) + \text{Var}(Z) = nP_Y(1 - P_Y) + nP_Z(1 - P_Z)$. The characteristic function is

$$(P_Y e^{it} + 1 - P_Y)^n (P_Z e^{it} + 1 - P_Z)^n$$

But unless $P_Y = P_Z$, there is no special name for the distribution of X .

EDIT: Maple does come up with a closed form for the probability mass function involving the associated Legendre function of the first kind:

$$\begin{aligned} \mathbb{P}(X = x) &= \\ &\begin{cases} \frac{n!}{x!} P_n^{x-n} \left(\frac{2P_Y P_Z - P_Y - P_Z}{P_Y - P_Z} \right) (P_Z - P_Y)^n \left(\frac{(1 - P_Z)(1 - P_Y)}{P_Z P_Y} \right)^{(n-x)/2} & \text{if } 0 \leq x \leq n \\ \frac{n!}{(2n - x)!} P_n^{n-x} \left(\frac{2P_Y P_Z - P_Y - P_Z}{P_Y - P_Z} \right) (P_Z - P_Y)^n \left(\frac{(1 - P_Z)(1 - P_Y)}{P_Z P_Y} \right)^{(n-x)/2} & \text{if } n \leq x \leq 2n \end{cases} \end{aligned}$$

EDIT: In response to Shakil's request, here is the Maple code:

```
> sum(binomial(n,k)*P[Z]^k*(1-P[Z])^(n-k)*
    binomial(n,x-k)*P[Y]^(x-k)*(1-P[Y])^(n-(x-k)),k=0..x) assuming x>=0,x<=n;
> simplify(%);
> sum(binomial(n,k)*P[Z]^k*(1-P[Z])^(n-k)*
    binomial(n,x-k)*P[Y]^(x-k)*(1-P[Y])^(n-(x-k)),k=x-n..n) assuming x>=n,x<=2*n;
> simplify(%);
```

edited Jul 28 at 22:00

answered Feb 17 '15 at 21:13



[Robert Israel](#)

227k 13 152 348

It has a special name: Poisson Binomial distribution. – [M. Alaggan](#) Feb 24 at 21:32

@Robert would you mind showing the maple code here. I do not know the maple at all. Thank you appreciate it. – [Shakil](#) Jul 28 at 19:04

See the [binomial sum variance inequality](#). Here is an excerpt from the Wikipedia page.

In probability theory and statistics, the sum of independent binomial random variables is itself a binomial random variable if all the component variables share the same success probability. If success probabilities differ, the probability distribution of the sum is not binomial.

answered Feb 17 '15 at 21:17



[Timothy Shields](#)

435 2 6

It will be a special case of the [Poisson Binomial Distribution](#).

answered Feb 24 at 21:32



[M. Alaggan](#)

335 2 16

In the limit as $n \rightarrow \infty$, your binomials become Gaussian and since it seems you are implicitly assuming your two binomials are independent, a sum of two independent Gaussians is Gaussian with mean and variance parameters given by the sum of the parameters for the two Gaussians, so yes, in the limit as $n \rightarrow \infty$, your distribution will converge to Binomial $(2n, (P_A + P_B)/2)$.

Also you are right, if $P_A = P_B = P$ and you assume independence, then the distribution is precisely Binomial $(2n, P)$.

However if $P_A \neq P_B$ and you assume independence, then the exact distribution is different from Binomial $(2n, (P_A + P_B)/2)$. If you let $X = X_A + X_B$ be the random variable which is the sum of your two binomials, then $P(X = k)$ is the summation over all the ways that you get $X_A = k_A$ and $X_B = k_B$ where $k_A + k_B = k$. It is easy to write down this summation formula if you know the formulas for binomial distribution, and summation notation. However I'm inclined to believe there is no closed form formula for it, unless it's something crazy like hypergeometric.

answered Feb 17 '15 at 21:08

[user2566092](#)



20.2k 16 38

"In the limit as $n \rightarrow \infty$, your binomials become Gaussian" Sorry but this is simultaneously vague and wrong. – [Did](#) Feb 17 '15 at 22:18

@Did You're a smart guy, you must know it can't be both vague and wrong at the same time. And in a certain sense, binomial does converge to Gaussian. Maybe I'm wrong in the conclusions about what that particular convergence means though. Maybe continuity corrections etc could save my argument, I don't know. But I thought that convergence to a Gaussian was enough to say what I was saying. – [user2566092](#) Feb 17 '15 at 23:10

And by Gaussian I mean normal distribution – [user2566092](#) Feb 17 '15 at 23:15

Vague: "binomials become Gaussian", meaning what? Wrong: Binomial (n, p) distributions do not converge to anything, gaussian or not, when $n \rightarrow \infty$. – [Did](#) Feb 17 '15 at 23:33

- 1 The De Moivre-Laplace theorem says that certain **scaled and translated** binomials converge to Gaussians. The binomials themselves do not converge (except in the trivial case of $p = 0$). – [Robert Israel](#) Feb 18 '15 at 1:59

|