Negative binomial distribution - sum of two random variables

Suppose X, Y are independent random variables with $X \sim NB(r,p)$ and $Y \sim NB(s,p)$. Then

$$X+Y \sim NB(r+s,p)$$

How do I go about proving this? I'm not sure where to begin, I'd be glad for any hint.

(probability) (statistics) (probability-distributions)

asked Dec 6 '14 at 8:16

iwriteonbananas
2,560 10 21

5 Answers

Hint:

If $\Pr(X=k) = \binom{k+r-1}{k} \cdot (1-p)^r p^k$ and $\Pr(Y=k) = \binom{k+s-1}{k} \cdot (1-p)^s p^k$ then

$$\Pr(X+Y=k) = \sum_{j=0}^{k} \binom{j+r-1}{j} \cdot (1-p)^r p^j \cdot \binom{k-j+s-1}{k-j} \cdot (1-p)^s p^{k-j}$$

$$= \sum_{j=0}^{k} \binom{j+r-1}{j} \cdot \binom{k-j+s-1}{k-j} \cdot (1-p)^{r+s} p^k$$

and you need to show

$$\Pr(X+Y=k) = \binom{k+r+s-1}{k} \cdot (1-p)^{r+s} p^k$$

so it is just a matter of showing
$$\sum_{j=0}^{k} \binom{j+r-1}{j} \cdot \binom{k-j+s-1}{k-j} = \binom{k+r+s-1}{k}.$$

answered Dec 6 '14 at 11:40

Henry
73.8k 3 44 108

Thank you very much! I don't understand why we have

$$\Pr(X+Y=k) = \sum_{j=0}^{k} \binom{j+r-1}{j} \cdot (1-p)^r p^j \cdot \binom{k-j+s-1}{k-j} \cdot (1-p)^s p^{k-j}$$

Could you please elaborate on how you know this? – iwriteonbananas Dec 6 '14 at 13:55

---

1 $Pr(X+Y=k) = \sum_j \Pr(X=j, Y=k-j)$ while $\Pr(X=j, Y=k-j) = \Pr(X=j)\Pr(Y=k-j)$ because they are independent – Henry Dec 6 '14 at 16:19

Ahhhhh I see, thank u! – iwriteonbananas Dec 6 '14 at 16:23

This is much simpler, can be seen directly from the definition without any calculations at all! – kjetil b halvorsen Nov 2 at 8:56

The $NB(r,p)$ can be written as independent sum of geometric random variables.

Let X_i be i.i.d. and $X_i \sim Geometric(p)$.

Then $X \sim NB(r,p)$ satisfies $X = X_1 + \cdots + X_r$,

and $Y \sim NB(s,p)$ satisfies $Y = X_{r+1} + \cdots + X_{r+s}$.

Therefore, $X + Y = X_1 + \cdots + X_{r+s}$.

This yields $X + Y \sim NB(r + s, p)$.

answered Dec 6 '14 at 9:06

i707107
7,103 1 8 26

Thank you! How can I see what the $NB(r, p)$ can be written as an independent sum of geometric random variables? – iwriteonbananas Dec 6 '14 at 9:14

If $X \sim NB(r, p)$, then $X = k$ means k is the time of r-th success. The geometric random variable gives the first time of success. – i707107 Dec 6 '14 at 9:16

I see, thank you. Where can I find a formal proof of this fact? – iwriteonbananas Dec 6 '14 at 9:19

Building upon the idea that NB(r,p) is the time to the r-th success in Bernoulli trials, and that the trials are independent, it is clear that NB(r+k,p) can be seen as the time to the r-th success and then to the next k-th success, giving the result directly with no algebra.

answered Jul 31 at 13:40

Naomi
21 1

This is the best answer, and deserves more upvotes! – kjetil b halvorsen Nov 2 at 8:56

This is essentially what I said in my answer. – i707107 Nov 2 at 20:39

Since X, Y are independent, the moment generating function (MGF) of $X + Y$ is the multiplication of the MGF of X and MGF of Y. The MGF of X is $M_X(t) = (\dfrac{1-p}{1-pe^t})^r$, and this is $(\dfrac{1-p}{1-pe^t})^s$ for Y. Now since X, Y are independent, we have that

$$\begin{aligned} M_{X+Y}(t) &= M_X(t)M_Y(t) \\ &= (\frac{1-p}{1-pe^t})^s(\frac{1-p}{1-pe^t})^r \\ &= (\frac{1-p}{1-pe^t})^{s+r} \end{aligned}$$

Therefore $M_{X+Y}(t) = (\frac{1-p}{1-pe^t})^{s+r}$ is the MGF of an NB distribution with parameters $r+s$ and p, meaning that $X+Y$ is $NB(r+s,p)$.

answered Dec 6 '14 at 8:29

Math-fun
5,622 8 22

Thank you very much! But I haven't seen the MGF in my course yet, and I'm wondering how to prove it without the use of MGF? – iwriteonbananas Dec 6 '14 at 8:31

Without MGF, you could think of the nature of an NB distribution and make intuitive arguments. – Math-fun Dec 6 '14 at 8:52

Have you learnt about the convolution of two independent random variables? That will allow you to compute the pmf directly without saying anything about the mgf. The method is to condition on one of them and use the total probability. For any $k \geq 0$, verify the sum is a NB pmf as required:

$$P(X+Y=k) = \sum_{x=0}^{k} P(Y+X=k|X=x)P(X=x) = \sum_{x=1}^{k} P(Y=k-x)P(X=x)$$

answered Dec 6 '14 at 8:57

QQQ
421 4 5