Matched Dispatching in Randomized Settings

BY Jostein Lillestøl and Per Manne

DISCUSSION PAPER

Institutt for foretaksøkonomi
Department of Business and Management Science

FOR 23/2023
ISSN: 2387-3000
December 2023

Matched Dispatching in Randomized Settings

by

Jostein Lillestøl and Per Manne

Department of Business and Management Science

Norwegian School of Economics, 5045 Bergen, Norway

November 29, 2023

Abstract

This paper examines some problems of matched dispatching in some different random settings. The context of presentation is that of a reality show with a lineup of the participants, and according to some probabilistic selection rule, some participants are pairwise matched to teams while some are excluded. We consider mainly two cases of induced randomness, one based on random ordering of participants, and one based on coin-flipping, and consider both linear and circular lineups. Questions of fairness are discussed, and some alternative schemes are examined.

JEL classification: C44, C46

Keywords: Random order, Matching, Dispatching, Hypergeometric distribution

1. Introduction

Patterns in sequences of observations are studied in many contexts, ranging from coin tossing for demonstrating the basics of probability to medical records for uncovering anomalies. Elementary probability asks questions related to the number of heads (H) in n trials and the waiting time to the first tail (T). The distribution and its expectation are typically established. Stepping up the ladder, one may ask for the number of successive patterns of a given type, say head-tail (HT). This becomes a bit harder, and typically belongs to intermediate texts, like Ross (2014).[1] Here follows another pattern problem with interesting features and challenges. Although a general problem, it is presented in a recreational context as follows:

[1] For more advanced combinatorics in probabilistic settings, the classical reference is Feller (1961).

In a TV reality show, celebrities are competing as recruits. At the beginning there are 14 participants, 7 men and 7 women. After some weeks 5 of them are out, and among the remaining 9, there are 5 men and 4 women. They are told to line up in random order facing the same direction. Then an officer commands: "Men: Turn right! Women: Turn left!". The result is that some recruits may face another recruit of the opposite gender, while some may face the back of a recruit of the same gender or no one at all. All pairs standing face to face are dispatched to act as mission teams, while the others are not teamed up, but ordered to do a dirty job, like renovating the latrine.

Questions to be asked are: What is the probability distribution of the number of teams and its expectation? What is expected number of recruits available for the dirty job? What can we conclude in general?

The reader may question the possibility to get people lined up in random order. To counter this legitimate objection, another context using playing cards may be imagined. However, for this presentation, we stick to the more frivolous one.

In case of 5 men and 4 women, a possible pattern for the 9 individuals is

M W M M W W M M W

Here we imagine the observer standing behind the lineup, facing the same direction. Given the command, a team will materialize whenever an M is followed by a W. We see that this occurs three times in the given pattern, so that 6 participants will be dispatched to teams, while 3 are not, in this case the males in the third and seventh position and the female in the sixth position. This is just one of 126 different equally likely patterns in this case. In order to calculate probabilities, we have to count the number of favorable patterns to the events in question. For instance, there are 40 different patterns that give rise to three teams. Consequently, the probability of this event is 40/126, by the rule "favorable on possible". The determination of the number of patterns exhibiting a given number of teams x apparently requires a complicated enumeration, unless we can establish a general formula.

In general, consider n participants, n_1 men and n_2 women. Then there will be $m = \binom{n_1+n_2}{n_1}$ different patterns of men and women all equally likely. Let X be the number of mission teams that materialize from the lineup after the command is given. The probability distribution of X is given by

$$P(X = x) = \frac{N(x)}{m}, \qquad x = 0, 1, \dots, \min(n_1, n_2)$$

where $N(x)$ is the number of lineups giving rise to x teams. Consider two simple examples:

Example 1: Take $(n_1, n_2) = (3, 2)$, for which we have $m = \binom{5}{3} = 10$ possible patterns:

WWMMM	MWWMM	WMWMM	WMMWM	WMMMW
MMWWM	MMMWW	MWMWM	MWMMW	MMWMW

Here MW does not appear in the first pattern, appears once in the next six patterns, and twice in the last three patterns. Consequently, we have the table

x	0	1	2
$N(x)$	1	6	3

Example 2: Take $(n_1, n_2) = (4, 3)$, for which we have $m = \binom{7}{4} = 35$ possible patterns, too many to list them all. The enumeration performed with programming help turned out the table

x	0	1	2	3
$N(x)$	1	12	18	4

The results for the two examples are seen to conform with the following general formula, to be proved in the subsequent section:

$$N(x) = \binom{n_1}{x} \cdot \binom{n_2}{x}, \qquad x = 0, 1, \ldots, \min(n_1, n_2).$$

The reader may work out and confirm the case $(n_1, n_2) = (3, 3)$, for which the distribution is symmetric.

Remark. We see that this expression is the same as the number of ways to select x men and x women to participation in one team or another. Does this mean that we have found a shortcut to the solution? No! Sample of persons should not be mistaken as our sample of positions in the lineup!

2. General theory

2.1 Random order: Linear lineup

Theorem 1: Given (n_1, n_2) and random linear lineup. Then the probability distribution of the number of teams X is given by

$$P(X = x) = \frac{\binom{n_1}{x} \cdot \binom{n_2}{x}}{\binom{n_1+n_2}{n_1}}, \qquad x = 0, 1, \ldots, \min(n_1, n_2),$$

recognized as the hypergeometric distribution with parameters $(n_1 + n_2, n_1, n_2)$ with expectation

$$E(X) = \frac{n_1 n_2}{n_1 + n_2}$$

Proof: Random lineup of the $n = n_1 + n_2$ participants means that all $n!$ possible lineups of the $n = n_1 + n_2$ participants are equally likely. What matters here is the pattern of consecutive genders. In total, there are $m = \binom{n_1 + n_2}{n_1}$ different possible patterns with n_1 men and n_2 women, all equally likely. This accounts for the denominator. In order to find the number of lineups having precisely x team, we may argue as follows: Line up the n_1 men in a row and choose the x positions of those to be included in one of the x teams. This may be done in $\binom{n_1}{x}$ ways. For every such selection we have to fit in the n_2 women so that one woman is immediately to the right of each of the x men assigned to a team. So let us do that. We are then left with $n_2 - x$ women to be fit in, without creating more teams. There are x+1 positions where these $n_2 - x$ women may fit in, either to the left of all men or immediately to the right of one of the x women already fitted in. To determine the number ways this can be done, we may use the following combinatorial result: Distribute a identical objects into b labelled boxes (allowing some empty boxes). This can be done in $\binom{a+b-1}{b-1}$ different ways. In our situation this corresponds to take $a = n_2 - x$ and $b = x + 1$, which inserted gives $\binom{n_2}{x}$ different ways. Taken together, this gives the number of lineups leading precisely to x teams equal to $N(x) = \binom{n_1}{x} \cdot \binom{n_2}{x}$. With this numerator we have proved the theorem.

Remark. The combinatorial argument used is often named "Stars and bars", after how it is typically explained: If the a objects are lined up in a row and marked with stars, the assignment to b boxes is by vertical bars as separation symbols. With b boxes we need $b - 1$ bars. Example: For $a = 4$ and $b = 3$ the pattern **|*|* means two objects in the first box and one object in each of the next two boxes, while **|**| means two objects in each of the first two boxes and none in the third box. All together there are $a + b - 1$ symbols, and the placement alternatives appear by choosing the $b - 1$ bars among them. For more on this theme see [Stars and bars (combinatorics) - Wikipedia] (Theorem 1 and Theorem 2).

The number of participants excluded from teams becomes $Y = n_1 + n_2 - 2X$ with expected value

$$E(Y) = n_1 + n_2 - 2E(X) = \frac{n_1^2 + n_2^2}{n_1 + n_2}$$

In general, the distribution is symmetric for $n_1 = n_2$, for which $E(X) = \frac{n}{4}$ and $E(Y) = \frac{n}{2}$ where $n = n_1 + n_2$. Thus, in the symmetric case, the expected number of participants assigned to teams will be the same as the expected number excluded.

Example 3: Take $(n_1, n_2) = (5, 4)$ and $(5, 5)$, with $m = \binom{9}{5} = 126$ and $m = \binom{10}{5} = 252$ possible patterns, respectively. Enumeration, performed with programming help and confirmed by the formula, turned out the table

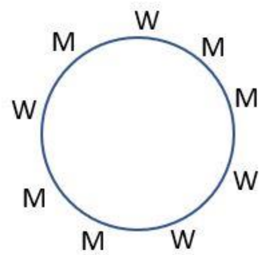
x	0	1	2	3	4	5	m
$N(x)$: $(5,4)$	1	20	60	40	5	–	126
$N(x)$: $(5,5)$	1	25	100	100	25	1	252

Expectations are: For $(n_1, n_2) = (5,4)$ we get $E(X)$ =20/9 and $E(Y) = 41/9$, and for $(n_1, n_2) = (5,5)$ we get $E(X) = 25/10 = 2.5$ and $E(Y) = 50/10 = 5$.

Note the asymmetry in the unbalanced case, and that the expected number of excluded participants 41/9 is just marginally larger than the expected number of participants assigned to teams 40/9.

2.2 Random order: Circular lineup

Consider the situation with a circular lineup. As an example, convert the following linear lineup WMMWWMMMWM to a circular one, starting on the top and going clockwise.



Imagine that the observer is inside the circle and the lineup is facing outward. We see that within the given linear lineup, there are only two instances of MW, while there are three instances in the circular lineup. Note that this circular pattern is identical to the one obtained from the linear nine-person lineup in Section 1, since the man at the left end of MWMMWWMMW has just moved over to the right end. This has given one less MW in the linear lineup, while there are still three in the circular lineup. Note also that we are sure to get at least one MW in a circular lineup.

Example 4: Take $(n_1, n_2) = (3,3)$ for which we have $m = 10$ possible patterns. In fact, they are given by putting a W up front for the ten linear patterns of Example 1 and tying the two ends together. The enumeration of the number of MW's now leads to the table

x	1	2	2
$N(x)$	3	6	1

Theorem 2: Given (n_1, n_2) and random circular lineup. Then the probability distribution of the number of teams X is given by

$$P(X = x) = \frac{\binom{n_1}{x} \cdot \binom{n_2 - 1}{x - 1}}{\binom{n_1 + n_2 - 1}{n_1 - 1}}, \qquad x = 1, 2, \ldots, \min{(n_1, n_2)}$$

This is recognized as the hypergeometric distribution with parameters $(n_1 + n_2 - 1, n_1, n_2)$ with expected value

$$E(X) = \frac{n_1 n_2}{n_1 + n_2 - 1}$$

Proof: In a circular lineup we imagine n marked locations around a circle with $n!$ different arrangements of the participants. However, the n arrangements obtained by rotation are equal from our point of view. We can therefore let one participant possess a fixed location, and let the others choose among the $n - 1$ remaining locations randomly. This can be done in $(n - 1)!$ different ways, all equally likely. Again, what matters is the pattern of males and females. Assume that the participant at the fixed location is a male. We then have $n_1 - 1$ males left to choose from $n - 1$ locations, leaving the rest to the females. This can be done in $\binom{n - 1}{n_1 - 1}$ different ways, all equally likely. This accounts for the denominator.

For the enumeration of the arrangements leading to x occurrences of MW, we may argue as in the proof of Theorem 1, using the combinatorial stars and bars formula. Again, we may first select the positions of the x men to enter a team among a lineup of the n_1 men. This can be done in $\binom{n_1}{x}$ ways. Then position the x female partners next to these males. The remaining $n_2 - x$ females have to be positioned so that no more MW's will occur. The only possibilities are next to an already assigned female. Circularity now rules out the ahead of all men opportunity we had in the linear case. There is no ahead of all, and any other placement will be double counting as well. Using the combinatorial formula from the proof of Theorem 1 with $a = n_2 - x$ and $b = x$ gives $\binom{n_2 - 1}{x - 1}$ possibilities for the females. In combination, we therefore have $\binom{n_1}{x} \cdot \binom{n_2 - 1}{x - 1}$ possibilities in all, which proves the numerator of the theorem. An alternative enumeration that may provide additional insight is given in Section 2.5.

The number of participants excluded from teams, $Y = n_1 + n_2 - 2X$, has expected value

$$E(Y) = n_1 + n_2 - 2E(X) = \frac{n_1(n_1 - 1) + n_2(n_2 - 1)}{n_1 + n_2 - 1}$$

In general, the distribution will be symmetric for $n_1 = n_2 + 1$ or $n_1 = n_2 - 1$.

Example 5: Take $(n_1, n_2) = (5, 4)$ and $(5, 5)$, for which we respectively have $m = \binom{8}{4} = 56$ and $m = \binom{9}{5} = 126$ possible patterns. Enumeration by the formula in Theorem 2 gave the table

x	1	2	3	4	5	m
$N(x)$: $(5, 4)$	4	24	24	4	-	56
$N(x)$: $(5, 5)$	5	40	60	20	1	126

Expectations are: For $(n_1, n_2) = (5, 4)$ we get $E(X) = \frac{20}{8} = 2.5$ and $E(Y) = 5 + 4 - 2 \cdot 2.5 = 4.0$, and for $(n_1, n_2) = (5, 5)$ we get $E(X) = \frac{25}{9}$ and $E(Y) = 5 + 5 - 2 \cdot \frac{25}{9} = \frac{40}{9}$. Not surprisingly, we expect a larger portion assigned to teams with the circular lineup than the linear one.

2.3 Coin flipping: Linear lineup

Consider n participants not identified by gender or other means. Again, consider a linear lineup with participants facing in the same direction. Then the following order is given: "Each one of you, flip a coin. If you got head (H) turn right. If you got tail (T) turn left". Then follows the same procedure as above: Two standing face to face to another will be a team, while the others remain unassigned to a team. We ask the same question: What is the probability distribution of the number of teams?

Different approached are available to analyze this problem. Here we take the opportunity to utilize the results obtained for the random order setup. This is possible by conditioning.

Theorem 3: Given n participants in a linear lineup with status Right turn or Left turn determined by individual coin flips. Then the probability distribution of the number of teams X is given by

$$P(X = x) = \binom{n+1}{2x+1} \cdot \left(\frac{1}{2}\right)^n , \qquad x = 0, 1, \dots, \left[\frac{n}{2}\right]$$

where $[x]$ is the truncated integer part of x. The expectation is $E(X) = \frac{n-1}{4}$.

Proof: Let N_1 be the number of heads in the n binomial trials. Conditionally, given $N_1 = n_1$, we are back to the random order setup, where M is replaced by H and W by T. Consequently,

$$P(X = x \mid N_1 = n_1) = \frac{\binom{n_1}{x} \cdot \binom{n - n_1}{x}}{\binom{n}{n_1}}$$

$$P(N_1 = n_1) = \binom{n}{n_1} \cdot \left(\frac{1}{2}\right)^n$$

Unconditionally we obtain

$$\begin{aligned} P(X = x) &= \sum_{n_1=0}^{n} P(X = x \mid N_1 = n_1) \cdot P(N_1 = n_1) \\ &= \sum_{n_1=0}^{n} \frac{\binom{n_1}{x} \cdot \binom{n - n_1}{x}}{\binom{n}{n_1}} \cdot \binom{n}{n_1} \cdot \left(\frac{1}{2}\right)^n \\ &= \left(\frac{1}{2}\right)^n \sum_{n_1=0}^{n} \binom{n_1}{x} \cdot \binom{n - n_1}{x} = \binom{n+1}{2x+1} \cdot \left(\frac{1}{2}\right)^n \end{aligned}$$

The last step follows from a binomial identity, the first of three binomial identities proven in the appendix. The formulas are valid for $x = 0, 1, \dots, [n/2]$, under the common conventions for binomial coefficients, that is $\binom{n}{0} = 1$ for $n \geq 0$ and $\binom{n}{x} = 0$ for $0 \leq n < x$ and for negative x. This will handle the case of $n_1 = 0$ and $n_1 = n$ as well.

The expectation may be found by conditioning as well

$$E(X) = E\big(E(X|N_1)\big) = E\left(\frac{N_1(n - N_1)}{n}\right) = \frac{n-1}{4}$$

The last step follows from the binomial properties $E(N_1) = \frac{n}{2}$ and $var(N_1) = \frac{n}{4}$.

This proves the theorem.

We may interpret the result within an outcome space of $m = 2^n$ possible outcomes of heads (H) and tails (T). We see that the solution has the form

$$P(X = x) = \frac{N(x)}{m} \text{ where } m = 2^n \text{ and } N(x) = \binom{n+1}{2x+1}$$

This corresponds to the set-up for the more direct combinatorial approach.

Remark. In our proof, the last sum expression is simply the number of ways we can get x pairs HT originating from the different partitions of the n participants and where you have to select x from the set of H's and x from the set of T's (one set possibly empty). However, this is a consequence and cannot be used as combinatorial proof. On the other hand, given a valid combinatorial proof, we have implicitly a proof of the sum identity.

Example 6: Calculations using the formulas of Theorem 3 give for the cases $n = 9, 10$ the table

x	0	1	2	3	4	5	m
$N(x): n = 9$	10	120	252	120	10	–	512
$N(x): n = 10$	11	165	462	330	55	1	1024

The expectations are $E(X) = \frac{9-1}{4} = 2$ and $E(X) = \frac{10-1}{4} = 2.25$, respectively.

In general, the distribution will be symmetric for n odd.

We may also embed the situation into an infinite string of participants. We then have the following corollary to Theorem 3.

Corollary: Let T_1 be the number in line of the first person that completes a team, i.e., the first occurrence of HT. The probability distribution is

$$P(T_1 = n) = (n-1) \cdot \left(\frac{1}{2}\right)^n, \qquad n = 2, 3, \ldots$$

Proof: Let $P_n = P(X = 0)$ for n given. From the theorem, $P_n = (n+1) \cdot \left(\frac{1}{2}\right)^n$. Then

$$P(T_1 = n) = P_{n-1} - P_n = n \cdot \left(\frac{1}{2}\right)^{n-1} - (n+1) \cdot \left(\frac{1}{2}\right)^n = (n-1) \cdot \left(\frac{1}{2}\right)^n$$

This result can of course be obtained by direct arguments. This, and the fact that $E(T_1) = 4$, is widely known.

2.4 Coin flipping: Circular lineup

Consider n participants not identified by gender or other means. Now, consider the circular lineup with participants facing outward from the center. As above, the following command is given: “Each one of you, flip a coin. If you get head (H), turn right. If you get tail (T), turn left”. Then follows the same procedure as above: Two participants standing face to face to another will be a team, while the others remain unassigned to a team. The probability distribution of the number of teams X may be derived by conditioning, using the results of Section 2.2. We now have

Theorem 4: Given n participants in a circular lineup receiving order to right turn or left turn determined by their individual coin flip. Then the probability distribution of the number of teams X is given by

$$P(X = x) = \binom{n}{2x} \cdot \left(\frac{1}{2}\right)^{n-1}, \qquad x = 0, 1, \dots, \left[\frac{n}{2}\right]$$

with expectation $E(X) = \frac{n}{4}$.

Proof: Let N_1 be the number of heads in the n binomial trials. Conditionally, given $N_1 = n_1$, we are back to the random order setup, where M is replaced by H and W by T. Consequently,

$$P(X = x \mid N_1 = n_1) = \frac{\binom{n_1}{x} \cdot \binom{n - n_1 - 1}{x - 1}}{\binom{n-1}{n_1 - 1}}$$

$$P(N_1 = n_1) = \binom{n}{n_1} \cdot \left(\frac{1}{2}\right)^n$$

Unconditionally, we obtain for $x = 1, 2, \dots, [n/2]$

$$\begin{aligned}
P(X = x) &= \sum_{n_1=0}^{n} P(X = x \mid N_1 = n_1) \cdot P(N_1 = n_1) \\
&= \sum_{n_1=1}^{n-1} \frac{\binom{n_1}{x} \cdot \binom{n - n_1 - 1}{x - 1}}{\binom{n-1}{n_1 - 1}} \cdot \binom{n}{n_1} \cdot \left(\frac{1}{2}\right)^n \\
&= \left(\frac{1}{2}\right)^n \sum_{n_1=1}^{n-1} \binom{n_1}{x} \cdot \binom{n - n_1 - 1}{x - 1} \cdot \frac{n}{n_1} = \left(\frac{1}{2}\right)^n \binom{n}{2x} \cdot 2
\end{aligned}$$

The last step follows from Binomial identity 2 in the Appendix, valid for $x > 0$, under the common conventions on binomial coefficients restated there. The terms for $n_1 = 0$ and $n_1 = n$ vanish, but reappear for $x = 0$, which fits the end formula as well.

The expectation obtained by conditioning and use of the expectation from Section 2.2 becomes

$$E(X) = E\big(E(X \mid N_1)\big) = E\left(\frac{N_1(n - N_1)}{n - 1}\right) = \frac{n}{4}$$

The last step follows from the binomial properties $E(N_1) = \frac{n}{2}$ and $var(N_1) = \frac{n}{4}$.

In general, the distribution will be symmetric for n even.

Example 7: Calculations using the formulas of Theorem 4 give for the cases $n = 9, 10$ the table

x	0	1	2	3	4	5	m
$N(x): n = 9$	2	12	252	168	18	–	512
$N(x): n = 10$	2	90	420	420	90	2	1024

The expectations are $E(X) = \frac{9}{4} = 2.25$ and $E(X) = \frac{10}{4} = 2.5$ respectively, that is slightly larger than in Example 6.

2.5 Conditional distributions

We consider again the (n_1, n_2) random linear lineup from Section 2.1 and examine some conditional probabilities that give additional insight to the scheme. Specifically, we look at conditional probabilities given the gender of the one taking the leftmost position. Let LM and LW denote the events that this person is respectively a male or a female. The conditional distributions are given by

$$P(X = x \mid LM) = \frac{P(X=x \cap LM)}{P(LM)} \quad \text{and} \quad P(X = x \mid LW) = \frac{P(X=x \cap LW)}{P(LW)}$$

Here the denominators are $P(LM) = \frac{n_1}{n_1+n_2}$ and $P(LW) = \frac{n_2}{n_1+n_2}$.

Theorem 5: Given (n_1, n_2) and random linear lineup. Then the conditional probability distributions of the number of teams X, given the gender of the leftmost person, are given by

$$P(X = x \mid LW) = \frac{\binom{n_1}{x} \cdot \binom{n_2 - 1}{x}}{\binom{n_1+n_2 - 1}{n_1}}, \qquad x = 0, 1, \dots, \min(n_1, n_2 - 1)$$

with expectation $E(X \mid LW) = \frac{n_1(n_2-1)}{n_1+n_2-1}$, and

$$P(X = x \mid LM) = \frac{\binom{n_1}{x} \cdot \binom{n_2 - 1}{x - 1}}{\binom{n_1+n_2 - 1}{n_1 - 1}}, \qquad x = 1, 2, \dots, \min(n_1, n_2)$$

with expectation $E(X \mid LM) = \frac{n_1 n_2}{n_1+n_2-1}$.

The first formula tells that given a leftmost woman, we are in the same situation as initially, but with one woman less to make it to team. This is as expected. The second formula is more involved, since given a leftmost man, this person may or may not be part of a team.

Proof: We address the conditioning on LM, where it may be instructive to adopt a more direct enumeration. By Theorem 1, there are $\binom{n_1-1}{x-1}\cdot\binom{n_2-1}{x-1}$ cases that begin with MW and has $x-1$ more teams following the first one. Again, by Theorem 1, there are $\binom{n_1-1}{x}\cdot\binom{n_2}{x}$ cases that begin with M and has x occurrences of MW among remaining n_1+n_2-1 persons. This will cover all the cases beginning with MM and containing in all x teams, but also the cases beginning with MW and containing x teams among the remaining n_1+n_2-2 persons. So we subtract the cases with $x+1$ teams in all. Simplifying by use of the Pascal triangle identity $\binom{n-1}{x-1}+\binom{n-1}{x}=\binom{n}{x}$ yields

$$\begin{aligned} N(x) &= \binom{n_1-1}{x-1}\cdot\binom{n_2-1}{x-1}+\binom{n_1-1}{x}\cdot\binom{n_2}{x}-\binom{n_1-1}{x}\cdot\binom{n_2-1}{x} \\ &= \binom{n_1-1}{x-1}\cdot\binom{n_2-1}{x-1}+\binom{n_1-1}{x}\cdot\binom{n_2-1}{x-1} \\ &= \binom{n_1}{x}\cdot\binom{n_2-1}{x-1} \end{aligned}$$

Hence we have obtained

$$P(X=x\cap LM)=\frac{\binom{n_1}{x}\cdot\binom{n_2-1}{x-1}}{\binom{n_1+n_2}{n_1}}$$

Consequently,

$$P(X=x\mid LM)=\frac{\binom{n_1}{x}\cdot\binom{n_2-1}{x-1}}{\binom{n_1+n_2}{n_1}}\cdot\frac{1}{\frac{n_1}{n_1+n_2}}=\frac{\binom{n_1}{x}\cdot\binom{n_2-1}{x-1}}{\binom{n_1+n_2-1}{n_1-1}}$$

The formulas are valid under the common conventions on binomial coefficients restated in the appendix.

Example 8: Consider the lineup in case of $(n_1,n_2)=(5,5)$. The unconditional and conditional probabilities $N(x)/m$ are determined by the table

x	0	1	2	3	4	5	m
$N(x)$	1	25	100	100	25	1	252
$LW: N(x)$	1	20	60	40	5	0	126
$LM: N(x)$	0	5	40	60	20	1	126

We see the distribution is shifted downwards for LW and upwards for LM. The three expectations are $E(X)=\frac{25}{10}=2.50$, $E(X\mid LW)=\frac{20}{9}=2.22$ and $E(X\mid LM)=\frac{25}{9}=2.77$.

3 Linear lineup with fixed ends

3.1 One end fixed: Random order

It may be argued that the linear lineup in Section 2.1 is not fair since a man at the right end and a woman at the left end have no opportunity to be selected to team at all. This is remedied by adopting the circular lineup. An alternative linear lineup would be to assign a man at the leftmost spot and a woman at the rightmost spot and have only the between ones randomized. In the context of the reality show this may be presented as a favor given to the winners within each gender of a preliminary challenge. We name this lineup $(1+n_1, 1+n_2)$. We will also consider lineups with one fixed and one open end, named $(1+n_1, n_2)$ and $(n_1, 1+n_2)$, indicating the gender of the added participant assigned to its favorable position. We may imagine a context where the loosing gender of a prior challenge faces the risk of getting the unfavorable spot. Consider first the $(1+n_1, n_2)$-lineup, for which we have

Theorem 6: Given the $(1+n_1, n_2)$ linear lineup with fixed assignment of man in the leftmost position and randomized order of the other (n_1, n_2). Then the probability distribution of the number of teams X is given by

$$P(X=x) = \frac{\binom{n_1+1}{x} \cdot \binom{n_2-1}{x-1}}{\binom{n_1+n_2}{n_1}}, \qquad x = 1, 2, \dots, \min(n_1+1, n_2)$$

recognized as the hypergeometric distribution with parameters (n_1+n_2, n_1+1, n_2) with expected value

$$E(X) = \frac{(n_1+1) \cdot n_2}{n_1+n_2}$$

Similarly, for the $(n_1, 1+n_2)$ lineup we have hypergeometric distribution with parameters (n_1+n_2, n_1, n_2+1).

Proof: The number of favorable patterns for $X=x$ may be determined as follows: There are two ways to get x men assigned to a team, either by (i) leftmost MW and then a selection of position of $x-1$ men among the n_1 men distributed randomly, or by (ii) leftmost MM and then a selection of x men among the n_1 men distributed randomly. For (i) we have the situation of Theorem 1 with (n_1, n_2-1) and $x-1$ replacing x, giving $\binom{n_1}{x-1} \cdot \binom{n_2-1}{x-1}$ possibilities. For (ii) we may argue similarly to the proof of Theorem 1, now with a fixed M at the left end position. Let the n_1 other men be lined up to the right and select the x men to be assigned to teams. This can be done in $\binom{n_1}{x}$ different ways. For each of these we position x accompanying women immediately to the right of the x men, creating a MW pattern, and then position the remaining $n_2 - x$ women so that no more MW patterns materialize. In this case, the only possibility is immediately to the right of any of the x already assigned women. Referring to the proof of Theorem 1, this corresponds to take $a = n_2 - x$ and $b = x$ in the combinatorial formula $\binom{a+b-1}{b-1}$, which gives $\binom{n_2-1}{x-1}$ possibilities. The number of ways according to (ii) is therefore $\binom{n_1}{x}\cdot\binom{n_2-1}{x-1}$. Taken together this gives

$$\begin{aligned} N(x) &= \binom{n_1}{x-1}\cdot\binom{n_2-1}{x-1}+\binom{n_1}{x}\cdot\binom{n_2-1}{x-1} \\ &= \left(\binom{n_1}{x-1}+\binom{n_1}{x}\right)\cdot\binom{n_2-1}{x-1} \\ &= \binom{n_1+1}{x}\cdot\binom{n_2-1}{x-1} \end{aligned}$$

where the last step follows by the Pascal triangle identity.

Example 9: Consider the $(1+n_1,n_2)$ lineup with fixed assignment at one end in the cases of $(n_1,n_2)=(4,4),(4,5)$, and $(5,5)$. We then get the table

x	1	2	3	4	5	m
$N(x)$: $(4,4)$	5	30	30	5	–	70
$N(x)$: $(4,5)$	5	40	60	20	1	126
$N(x)$: $(5,5)$	6	60	120	60	6	252

Note that this distribution matches the conditional distribution, given a leftmost male, as derived in Section 2.5. The difference is only conceptual: Happened by chance versus fixed by choice. For comparisons: n_1 of Theorem 5 corresponds to $1+n_1$ in Theorem 6, following the convention that (n_1,n_2) represent the randomized participants. The case of $(n_1,n_2)=(4,5)$ here is therefore identical to $(n_1,n_2)=(5,5)$ of Example 8.

3.2 Both ends fixed: Random order

Consider the case of fixed assignment at both ends, avoiding both possibly unfavorable assignments. We then have

Theorem 7: Given a $(1+n_1,1+n_2)$ linear lineup with fixed assignments at both ends and randomized order of (n_1,n_2). Then the probability distribution of the number of teams X is given by

$$P(X=x)=\frac{\binom{n_1}{x-1}\cdot\binom{n_2}{x-1}}{\binom{n_1+n_2}{n_1}},\qquad x=1,2,\dots,\min(n_1+1,n_2+1)$$

recognized as a shifted hypergeometric distribution with parameters (n_1+n_2,n_1,n_2), having expected value

$$E(X) = 1 + \frac{n_1 \cdot n_2}{n_1 + n_2}$$

Proof: The number of favorable patterns for $X = x$ may be determined using counting principles as above. This time we have four cases according to whether we get none, one or two teams due to the fixe assignments at the ends of the lineup. This leads to

$$\begin{aligned} N(x) &= \binom{n_1 - 1}{x - 2} \cdot \binom{n_2 - 1}{x - 2} + \binom{n_1 - 1}{x - 1} \cdot \binom{n_2 - 1}{x - 2} + \binom{n_1 - 1}{x - 2} \cdot \binom{n_2 - 1}{x - 1} + \binom{n_1 - 1}{x - 1} \cdot \binom{n_2 - 1}{x - 1} \\ &= \left(\binom{n_1 - 1}{x - 1} + \binom{n_1 - 1}{x - 2}\right) \cdot \left(\binom{n_2 - 1}{x - 1} + \binom{n_2 - 1}{x - 2}\right) \\ &= \binom{n_1}{x - 1} \cdot \binom{n_2}{x - 1} \end{aligned}$$

where the simplification again comes from the Pascal triangle identity.

Example 10: Consider $(1 + n_1, 1 + n_2)$ lineup in case of $(n_1, n_2) = (4, 4), (4, 5)$ and $(5, 5)$. We get the table

x	1	2	3	4	5	6	m
$N(x): (4, 4)$	1	16	36	16	1	–	70
$N(x): (4, 5)$	1	20	60	40	5	–	126
$N(x): (5, 5)$	1	25	100	100	25	1	252

3.3 One end fixed: Coin flipping

Consider the linear lineup with the left end fixed as described in Section 3.1, but with the action determined by coin flipping as in Section 2.3. That is, participants are not identified by gender or other means and the command given is: “Each one of you, flip a coin. If you got head (H) turn right. If you got tail (T) turn left”. A special rule is added, forcing the person in fixed left position to take a right turn. Then, after the flipping, we face the result equivalent to a $(1 + n_1, n_2)$ scheme with $n = n_1 + n_2$, as described in Section 3.1. We now have

Theorem 8: Given n participants in a linear lineup with status Right turn or Left turn determined by individual coin flips with an additional participant at left forced to turn right. Then the probability distribution of the number of teams X is given by

$$P(X = x) = \binom{n + 1}{2x} \cdot \left(\frac{1}{2}\right)^n, \qquad x = 0, 1, \ldots, \left[\frac{n + 1}{2}\right]$$

with expectation $E(X) = \frac{n+1}{4}$.

Proof: Let N_1 be the number of heads in the n binomial trials. Conditionally, given $N_1 = n_1$, we are back to the random order setup, where M is replaced by H and W by T. Consequently,

$$P(X = x \mid N_1 = n_1) = \frac{\binom{n_1 + 1}{x} \cdot \binom{n - n_1 - 1}{x - 1}}{\binom{n}{n_1}}$$

$$P(N_1 = n_1) = \binom{n}{n_1} \cdot \left(\frac{1}{2}\right)^n$$

Unconditionally, we obtain

$$\begin{aligned} P(X = x) &= \sum_{n_1=0}^{n} P(X = x \mid N_1 = n_1) \cdot P(N_1 = n_1) \\ &= \sum_{n_1=0}^{n} \frac{\binom{n_1 + 1}{x} \cdot \binom{n - n_1 - 1}{x - 1}}{\binom{n}{n_1}} \cdot \binom{n}{n_1} \cdot \left(\frac{1}{2}\right)^n \\ &= \left(\frac{1}{2}\right)^n \sum_{n_1=0}^{n} \binom{n_1 + 1}{x} \cdot \binom{n - n_1 - 1}{x - 1} = \binom{n + 1}{2x} \cdot \left(\frac{1}{2}\right)^n \end{aligned}$$

Here we have used Binomial identity 3 in the Appendix. Again, the validity for $x = 0, 1, \dots, n$ is under the common conventions for binomial coefficients given in the Appendix.

The expectation may be found by conditioning as well,

$$E(X) = E\big(E(X|N_1)\big) = E\left(\frac{(N_1 + 1)(n - N_1)}{n}\right) = \frac{n + 1}{4}$$

The last step follows from the binomial properties $E(N_1) = \frac{n}{2}$ and $var(N_1) = \frac{n}{4}$.

This proves the theorem.

Example 11: Calculations using the formulas of Theorem 7 give for the cases $n = 9, 10$ the table

x	0	1	2	3	4	5	6	m
$N(x): n = 9$	1	45	210	210	45	1	–	512
$N(x): n = 10$	1	55	330	462	165	11	1	1024

3.4 Both ends fixed: Coin flipping

Consider the linear lineup with coin flipping as described in the preceding section, but with both ends fixed. Again, a special rule is adopted, forcing the persons in the fixed outside positions to take inward turns. Then, after the flipping, we face a situation equivalent to a $(1+n_1, 1+n_2)$ scheme with $n = n_1 + n_2$, as described in Section 3.2. We now have

Theorem 9: Given n participants in a linear lineup with status Right turn or Left turn determined by individual coin flips with additional end participants forced to turn inwards. Then the probability distribution of the number of teams X is given by

$$P(X = x) = \binom{n+1}{2x-1} \cdot \left(\frac{1}{2}\right)^n, \qquad x = 1, 2, \ldots, \left\lceil \frac{n+1}{2} \right\rceil$$

with expectation $E(X) = 1 + \frac{n-1.}{4}$.

Proof: Let N_1 be the number of heads in the n binomial trials. Conditionally, given $N_1 = n_1$, we are back to the random order setup, where M is replaced by H and W by T. Consequently,

$$P(X = x \mid N_1 = n_1) = \frac{\binom{n_1}{x-1} \cdot \binom{n-n_1}{x-1}}{\binom{n}{n_1}}$$

$$P(N_1 = n_1) = \binom{n}{n_1} \cdot \left(\frac{1}{2}\right)^n$$

Unconditionally, we obtain

$$\begin{aligned} P(X = x) &= \sum_{n_1=0}^{n} P(X = x \mid N_1 = n_1) \cdot P(N_1 = n_1) \\ &= \sum_{n_1=0}^{n} \frac{\binom{n_1}{x-1} \cdot \binom{n-n_1}{x-1}}{\binom{n}{n_1}} \cdot \binom{n}{n_1} \cdot \left(\frac{1}{2}\right)^n \\ &= \left(\frac{1}{2}\right)^n \sum_{n_1=0}^{n} \binom{n_1}{x-1} \cdot \binom{n-n_1}{x-1} = \binom{n+1}{2x-1} \cdot \left(\frac{1}{2}\right)^n \end{aligned}$$

Here we have used Binomial identity 1 in the Appendix with x replaced by $x-1$. Again, the validity is dependent on the common conventions for binomial coefficients.

The expectation may be found by conditioning

$$E(X) = E\big(E(X|N_1)\big) = E\left(1 + \frac{N_1 \cdot (n - N_1)}{n}\right) = 1 + \frac{n-1}{4}$$

Alternatively, just note the distribution shift and expectation in Theorem 3.

This proves the theorem.

Example 12: Calculations using the formulas of Theorem 8 give for the cases $n = 9, 10$ the table

x	0	1	2	3	4	5	6	m
$N(x): n = 9$	–	10	120	252	120	10	–	512
$N(x): n = 10$	–	11	165	462	330	55	1	1024

4 Some aspects of fairness

Above we have argued that the linear lineup in Section 2.1 may be felt unfair and have presented modifications in Section 3, pretending to fix the problem and, at the same time, opening up for favoring. We will now discuss aspects of fairness in relation to these schemes. In particular, we examine the individual probabilities of being assigned to a team.

The common interpretation of fairness would be that all participants have the same probability of being assigned to a team. In case one or more participants are given a special treatment at the outset, the regular ones should face the same probability. In this case, it may be interesting to determine the given advantage in probability terms.

4.1 Dispatch probabilities: Free ends

Consider the linear (n_1, n_2)-lineup with free ends, and define the events

$$M_j \ : \text{man no.} \, j \text{ is assigned to team}, \qquad j = 1, 2, \dots, n_1$$

$$W \ : \text{woman no.} \, j \text{ is assigned to team}, \qquad j = 1, 2, \dots, n_2$$

For a random order linear (n_1, n_2)-scheme with $n = n_1 + n_2$ we have

$$P(M_j) = \frac{(n-1) \cdot n_2 \cdot (n-2)!}{n!} = \frac{n_2}{n}, \qquad j = 1, 2, \dots, n_1$$

$$P(W_j) = \frac{(n-1) \cdot n_1 \cdot (n-1)!}{n!} = \frac{n_1}{n}, \qquad j = 1, 2, \dots, n_2$$

The argument: All $n!$ orderings are equally likely. Hence the denominator is $m = n!$. A given male may be assigned to team in one of $n-1$ positions having one of n_2 females to his right and the other $n-2$ participants in a random order, in total $(n-2)!$ different ones. The argument for the women is similar with n_1 replacing n_2 in the formulas.

This demonstrates that for the scheme to be fair to all participants, we must have $n_1 = n_2$. Otherwise, the gender with fewest participants is favored.

In case of $n_1 = n_2$, the linear scheme will be fair, in the sense that no one is favored. They all face the same probability of being assigned to team. In particular, they run the same risk of taking the unfavorable end position. As the scheme unfolds, such a participant may nevertheless feel cheated. In this respect, the circular lineup is preferable, as it avoids this kind of perceived unfairness. However, the gender unfairness for circular schemes will remain whenever $n_1 \neq n_2$.

Assume, after random ordering, that we observe who is at the left end position. Without loss of generality, we label this person 1 and let LM denote a man and LW a woman.

Conditional probabilities are given by

$$P(M_j \mid LM) = \frac{P(M_j \cap LM)}{P(LM)} = \frac{(n-2)\cdot n_2 \cdot (n-3)!/n!}{1/n} = \frac{n_2}{n-1}, \qquad j = 1, 2, \dots, n_1$$

$$P(W_j \mid LM) = \frac{P(W_j \cap LM)}{P(LM)} = \frac{(n-2)\cdot n_1 \cdot (n-3)!/n!}{1/n} = \frac{n_1}{n-1}, \qquad j = 1, 2, \dots, n_2$$

Note that all $P(M_j \mid LM)$ are equal, also for male participant no. 1 in the presumed favorable position. Note that the probabilities of being dispatched as team have increased for both males and females to the same degree. Both genders benefitted from a favorable left positioning that happened by chance. Similarly, we have

$$P(M_j \mid LW) = \frac{P(M_j \cap LW)}{P(LM)}$$

$$= \frac{(n-2)\cdot (n_2 - 1) \cdot (n-3)!/n!}{1/n} = \frac{n_2 - 1}{n-1}, \qquad j = 1, 2, \dots, n_1$$

$$P(W_j \mid LW) = \frac{P(W_j \cap LW)}{P(LM)}$$

$$= \frac{(n-2)\cdot (n_1 - 1) \cdot (n-3)!/n!}{1/n} = \frac{n_1 - 1}{n-1}, \qquad j = 1, 2, \dots, n_2$$

As expected, these dispatch probabilities are all reduced, due to the blocking of a favorable opportunity. Similar results follow by conditioning on the right end.

We may expect the same effect if we deliberate force this positioning at the outset. Before looking into this, we take the opportunity to show an alternative way of deriving the expected number of established teams: Let I_j be the indicator of the event M_j, that is $I_j = 1$ if M_j is true and $I_j = 0$ otherwise, for $j = 1, 2, \dots, n_1$. Then $X = I_1 + I_2 + \dots + I_{n_1}$ and we have a sum of terms with the same expectation $E(I_j) = P(I_j = 1) = \frac{n_2}{n}$. We therefore get

$$E(X) = E(I_1 + I_2 + \dots + I_{n_1}) = E(I_1) + E(I_2) + \dots + E(I_{n_1}) = n_1 \cdot \frac{n_2}{n}.$$

Alternatively, we may use indicators J_i for the female events W_i for $i = 1, 2, \dots, n_2$ and write $X = J_1 + J_2 + \dots + J_{n_2}$, so that we get $E(X) = n_2 \cdot \frac{n_1}{n}$. Derivation of the variance along this line is also possible, but has to take into account that indicators are correlated.

4.2 Dispatch probabilities: Fixed ends

Consider a $(1+n_1, n_2)$-lineup with a man labelled 0 fixed at the left end, that is in total $n = 1 + n_1 + n_2$ participants. Otherwise, the notation is as above. For a random ordering of the other $n_1 + n_2$ participants we have by arguing similarly to the above

$$P(M_j) = \frac{n_2}{n-1}, \qquad j = 0, 1, \dots, n_1$$

$$P(W_j) = \frac{n_1 + 1}{n-1}, \qquad j = 1, 2, \dots, n_2$$

Although the added participant 0 is freed from the risk of ending in the, for him, unfavorable rightmost position, he has no advantage in probability terms of being assigned to a team. We see that the scheme is fair within gender, and wholly fair only for $n_2 = n_1 + 1$, that is when the number of men and women in the lineup is equal. Otherwise, the less frequent gender is favorized. The number of teams may again be expressed by a sum of indicators, and its expectation confirmed as the one given in Section 3.1.

Then, consider a $(1+n_1, 1+n_2)$-lineup with a man labelled 0 fixed at the left end and a woman 0 at the right end, that is in total $n = 2 + n_1 + n_2$ participants. Otherwise, the notation is as above. For a random ordering of the other $n_1 + n_2$ participants we have by arguing similarly to the above

$$P(M_0) = \frac{n_2}{n-2}, \qquad P(W_0) = \frac{n_1}{n-2}$$

$$P(M_j) = \frac{n_2 + 1}{n-2}, \qquad j = 1, 2, \dots, n_1$$

$$P(W_j) = \frac{n_1 + 1}{n-2}, \qquad j = 1, 2, \dots, n_2$$

Now it turns out, possibly as a surprise, that the fixed end positions are disfavored compared to the free participants of the same gender. The number of teams may again be expressed by a sum of indicators, and its expectation confirmed as the one given in Section 3.1.

$$E(X) = E(I_0) + E(I_1) + \dots + E(I_{n_1}) = \frac{n_2}{n-2} + n_1 \cdot \frac{n_2 + 1}{n-2} = 1 + \frac{n_1 \cdot n_2}{n-2}$$

For comparison with the open-end situation, note again that n_1 and n_2 here should be reduced by 1.

4.3 Randomness

We must assume that none of the participants have knowledge about the command to follow. If there is knowledge, they may plan to line up according to their own preference. To safeguard, the commander may randomize the command, that is flip a coin to decide which gender to turn right and the other left.

With the knowledge later obtained they may regret that they did not line up alternately man/woman, and next to their favorite as well. This will of course be completely contrary to the assumed random order. Anyway, it may seem impossible in practice to arrange a random order based on individual behavior. Friends may stick together, and so on. An imposed randomization mechanism may be required here as well.

An alternative context where this problem does not arise is using playing cards, by letting the card color black and red represent each gender. We may then pick n_1 black cards to represent the men and n_2 red cards to represent the women. The pile of $n = n_1 + n_2$ cards is then shuffled, and cards are laid out from the top. We then look for the pattern BR.

5 Extensions

We have developed a theory for linear and circular lineups with a specific matching rule. Various extensions may be imagined in different directions, mainly

a. Different lineups

b. Different matching rules

c. More than two object categories

These extensions will, to varying degree, require more challenging mathematics.

For linear and circular lineups, we may imagine matching rules involving more than the two-letter pattern MW. We may also imagine situations where a disruptive person enters the lineup at a random position, and possibly spoils a match.

Next, we may imagine situations with more than two categories, and a matching rule adapted to this. In order to stay within the gender context, we may consider the following: Among the n participants there are n_0 who do not define themselves into the male/female dichotomy, so that $n = n_0 + n_1 + n_2$. If we want to keep the dichotomic matching rule, we have three ways of matching participants of the added gender group: Let them match with both common genders, or match with none of the common genders, or make matches according to a rule based on coin flips. More challenging will be to treat the situation as a genuine three-category problem, possibly with three-letter matching rules. The command given must of course be modified to account for the added group context.

Finally, we may imagine two-dimensional lineups, where participants are positioned randomly in a rectangular pattern. Different commands leading to matchings may be imagined, involving just one direction or more. A simple example is a $2 \times n$ lineup, where participants are queuing up two by two ready to march forwards, like we did in primary school. Suppose that the command is given as above, then affecting just neighboring pairs. It turns out that this problem may be solved using trinomial coefficients of a certain kind, see Andrews (1990).

Appendix: Three binomial identities

Identities for the sum of products of binomial coefficients with summation over the upper index are not commonplace, and we provide here three required formulas with proofs.

The following common conventions on binomial coefficients are used: $\binom{n}{0} = 1$ for $n \geq 0$ and $\binom{n}{x} = 0$ for $0 \leq n < x$ and for $x < 0$. Moreover, we take $\binom{n}{x} = 0$ for $n < x < 0$, while $\binom{n}{n} = 1$ for $n < 0$. Binomial coefficients for negative top indices are not that common, but can be defined consistently, so that basic relationships remain true, including the Pascal triangle identity, see Kronenburg (2015). This offers the opportunity to have summations going from 0 to n, and to extend validity of formulas.

Binomial identity 1:

$$\sum_{k=0}^{n} \binom{k}{x} \cdot \binom{n-k}{x} = \binom{n+1}{2x+1}, \qquad x = 0, 1, 2, \dots, \left\lceil \frac{n}{2} \right\rceil$$

Proof: Let $\{c_n; n = 0, 1, 2, \dots.\}$ be the sequence of sums, with generating function $C(z) = \sum_{n=0}^{\infty} c_n \cdot z^n$. This sum is recognized as the convolution $\{c_n\} = \{a_n\} * \{a_n\}$ of the sequence $a_n = \binom{n}{x}$, with generating function

$$A(z) = \sum_{n=0}^{\infty} a_n \cdot z^n = \sum_{n=0}^{\infty} \binom{n}{x} \cdot z^n = \frac{z^x}{(1-z)^{x+1}}$$

Consequently, the generating function of $\{c_n\}$ becomes

$$\begin{aligned} C(z) &= A(z) \cdot A(z) = \frac{z^x}{(1-z)^{x+1}} \cdot \frac{z^x}{(1-z)^{x+1}} \\ &= \frac{z^{2x}}{(1-z)^{2x+2}} = \frac{1}{z} \cdot \frac{z^{2x+1}}{(1-z)^{2x+1+1}} \end{aligned}$$

Here the second factor is the generating function of $\{a_n\}$ with x replaced by $2x + 1$. Consequently,

$$C(z) = \frac{1}{z} \sum_{n=1}^{\infty} \binom{n}{2x+1} \cdot z^n = \sum_{n=0}^{\infty} \binom{n+1}{2x+1} \cdot z^n$$

From this we read the desired result $c_n = \binom{n+1}{2x+1}$.

Binomial identity 2:

$$\sum_{k=1}^{n-1}\binom{k}{x}\cdot\binom{n-k-1}{x-1}\cdot\frac{n}{k}=\binom{n}{2x}\cdot 2,\qquad x=1,2,\dots,\left[\frac{n}{2}\right]$$

Proof: Write the sum as $n\ \cdot c_n$ where $c_n=\sum_{k=1}^{n}\frac{1}{k}\binom{k}{x}\cdot\binom{n-1-k}{x-1}$ is recognized as the convolution $\{c_n\}=\{a_n\}*\{b_n\}$ of the sequences $a_n=\frac{1}{n}\binom{n}{x}$ and $b_n=\binom{n-1}{x}$. Their generating functions are, respectively

$$A(z)=\sum_{n=1}^{\infty}a_n\cdot z^n=\sum_{n=1}^{\infty}\frac{1}{n}\cdot\binom{n}{x}\cdot z^n=\frac{1}{x}\cdot\frac{z^x}{(1-z)^{x+1}}$$

$$B(z)=\sum_{n=1}^{\infty}b_n\cdot z^n=\sum_{n=1}^{\infty}\binom{n-1}{x}\cdot z^n=\frac{z^x}{(1-z)^x}$$

The generating function of $\{c_n\}$ is therefore

$$\begin{aligned}C(z)&=A(z)\cdot B(z)=\frac{1}{x}\cdot\frac{z^x}{(1-z)^{x+1}}\cdot\frac{z^x}{(1-z)^x}\\&=\frac{1}{x}\cdot\frac{z^{2x}}{(1-z)^{2x+1}}=2\ \cdot\frac{1}{2x}\cdot\frac{z^{2x}}{(1-z)^{2x+1}}\end{aligned}$$

Except the factor 2, this is similar to $A(z)$ with x replaced by $2x$. We therefore have $c_n=\frac{2}{n}\binom{n}{2x}$ and consequently $n\cdot c_n=2\cdot\binom{n}{2x}$, which proves the identity.

Binomial identity 3:

$$\sum_{k=0}^{n}\binom{k+1}{x}\cdot\binom{n-k-1}{x-1}=\binom{n+1}{2x},\qquad x=1,2,\dots,\left[\frac{n}{2}\right]$$

Proof: The sequence of sum $\{c_n\}$ is recognized as a convolution $\{a_n\}*\{b_n\}$ of the sequences $a_n=\binom{n+1}{x}$ and $b_n=\binom{n-1}{x-1}$. Their generating functions are, respectively

$$A(z)=\sum_{n=0}^{\infty}a_n\cdot z^n=\sum_{n=0}^{\infty}\binom{n+1}{x}\cdot z^n=\frac{z^{x-1}}{(1-z)^{x+1}}$$

$$B(z)=\sum_{n=0}^{\infty}b_n\cdot z^n=\sum_{n=0}^{\infty}\binom{n-1}{x-1}\cdot z^n=\frac{z^x}{(1-z)^x}$$

The generating function of $\{c_n\}$ is therefore

$$\begin{aligned} C(z) &= A(z) \cdot B(z) \\ &= \frac{z^{x-1}}{(1-z)^{x+1}} \cdot \frac{z^x}{(1-z)^x} = \frac{z^{2x-1}}{(1-z)^{2x+1}} \end{aligned}$$

This is similar to $A(z)$ with x replaced by $2x$. We therefore have $c_n = \binom{n+1}{2x}$, which proves the identity.

Note: Attention is required to see that the case of $k = n$ with $x = 0$ fits in, using the conventions of binomial coefficients with negative arguments stated above.

It is worthwhile to note that the identities will appear as convolutions for diagonal sequences in Pascal's triangle. A more general formula of this kind, in terms of two variables, may be found in Feller (1961). To be specific, it is formula 12.16, given as theoretical problem in Section 12 of Chapter II. The formula may also be found as formula 3.2 of Gould (1972), again with no proof. A version of the formula with proof is given as formula (11) in § 9 of Netto (1901), possibly its origin. It turns out, by clever substitutions and rearrangements of terms, that our three formulas may be brought into the common framework of these formulas.

References

Andrews G. (1990) Euler's "Exemplum memorabile inductionis fallacis" and q-trinomial coefficients, J. Amer. Math. Soc., 3, 653–669.

Feller W. (1961) An Introduction to Probability Theory and its Applications, Vol. 1, 2. ed., New York: Wiley.

Gould H. W. (1972) Combinatorial identities: a standardized set of tables listing 500 binomial coefficient summation, Morgantown Rev. Ed., W. Va: Gould.

Kronenburg, M. J. (2015) The binomial coefficient for negative arguments, arXiv:1105.3689v2 **[math.CO]**, https://doi.org/10.48550/arXiv.1105.3689

Netto, E. (1901) Lehrbuch der Combinatorik, Leipzig: B. G. Teubner.

Ross S. M. (2014) Introduction to Probability Models, 11th ed., Oxford: Elsevier.

Wikipedia (2023) Stars and bars (combinatorics). Last visited Oct. 25, 2023.

NHH

NORGES HANDELSHØYSKOLE
Norwegian School of Economics

Helleveien 30
NO-5045 Bergen
Norway

T +47 55 95 90 00
E nhh.postmottak@nhh.no
W www.nhh.no