Solution to G-Research Sample Quant Exam Question 11

Mario Veca

31 January 2025

Introduction

This document provides a detailed solution to the final problem in the G-Research Sample Quant Exam, labeled as the hardest (and hence the most interesting). The question challenges players to select a positive integer in a way that maximizes their chances of winning, given that the smallest unique choice wins the game. Despite its simple rules, solving the problem requires creativity, careful reasoning, and a solid grasp of Nash equilibria.
This solution derives the Nash equilibrium equations, finds a solution and proves its uniqueness. The original problem can be found in the G-Research Sample Quant Exam [1].

The Problem

The statement of the problem can be enunciated as follows:

"You play a game with three players (including you), where each player chooses a positive integer. The player who picks the smallest number not chosen by anyone else wins a prize. If all the players choose the same number, nobody wins. How should you make your choice?"

The goal of this document is to derive a detailed solution to this problem, including:

1. Finding the Nash equilibrium strategy equations for each player.
2. Finding a solution to the equations using an ansatz.
3. Providing a proof of uniqueness for the ansatz solution.

Solution

Nash Equilibrium equations

This is a game theory problem for three player with countable infinite choices. Since it is clear that there is no dominant strategy, we need to find the Nash Equilibrium/a for the optimal mixed-strategies.

We define the following probabilities:

- p_n: the probability that the first player (Player A) chooses n,
- q_n: the probability that the second player (Player B) chooses n,

- r_n: the probability that the third player (Player C) chooses n.

The expected payoff for Player A when choosing the number n is:

$$\pi_A(n) = \sum_{k=n+1}^{\infty} \sum_{j=n+1}^{\infty} q_k r_j + \sum_{k=1}^{n-1} q_k r_k, \tag{1}$$

where the first term accounts for cases where n is unique, and the second term represents cases where n ties (the sum has to be considered 0 if the upper limit is less than the lower limit). Similarly, the expected payoffs for Players B and C are:

$$\pi_B(n) = \sum_{k=n+1}^{\infty} \sum_{j=n+1}^{\infty} p_k r_j + \sum_{k=1}^{n-1} p_k r_k, \tag{2}$$

$$\pi_C(n) = \sum_{k=n+1}^{\infty} \sum_{j=n+1}^{\infty} p_k q_j + \sum_{k=1}^{n-1} p_k q_k. \tag{3}$$

Then, the Nash equilibrium is achieved when, for each player, the expected payoffs associated with the various choices of n are all equal. This can be done by equating, for all n, the payoff for choosing n and for choosing $n+1$:

$$\pi_A(n) = \pi_A(n+1), \tag{4}$$

$$\pi_B(n) = \pi_B(n+1), \tag{5}$$

$$\pi_C(n) = \pi_C(n+1). \tag{6}$$

Writing explicitly and simplifying the equation for player A, we get,

$$q_{n+1} \sum_{k=n+1}^{\infty} r_k + r_{n+1} \sum_{k=n+1}^{\infty} q_k - q_{n+1} r_{n+1} = q_n r_n. \tag{7}$$

Similarly, we can derive analogous equations for Players B and C, finding the equation system,

$$q_{n+1} \sum_{k=n+1}^{\infty} r_k + r_{n+1} \sum_{k=n+1}^{\infty} q_k - q_{n+1} r_{n+1} = q_n r_n \ , \tag{8}$$

$$p_{n+1} \sum_{k=n+1}^{\infty} r_k + r_{n+1} \sum_{k=n+1}^{\infty} p_k - p_{n+1} r_{n+1} = p_n r_n \ , \tag{9}$$

$$q_{n+1} \sum_{k=n+1}^{\infty} p_k + p_{n+1} \sum_{k=n+1}^{\infty} q_k - p_{n+1} q_{n+1} = q_n p_n \ . \tag{10}$$

By symmetry, the solution of the system must[1] be have $p_k = q_k = r_k$ for all k. Therefore, we are left with the following simplified equation system:

$$2p_{n+1} \sum_{k=n+1}^{\infty} p_k - p_{n+1}^2 = p_n^2. \tag{11}$$

[1]While this is quite intuitive, we will provide a rigorous proof of this in the Appendix.

Solution Ansatz

We observe that in the above equation only polynomials of the same grade (two) appear, hence, if $\mathbf{P}^*$ satisfies the above equality, then $\alpha\mathbf{P}^*$ also satisfies it. This property can later be used to enforce the normalization condition $\sum_{k=1}^{\infty} p_k = 1$.
We can solve the above equation with the following exponential ansatz:

$$p_n = \alpha x^{n-1}, \tag{12}$$

where x is a parameter such that $0 < x < 1$, and α is a normalization constant.
Substituting the ansatz into (11), we find

$$\begin{aligned} &2x^n \sum_{k=n+1}^{\infty} x^{k-1} - x^{2n} = x^{2(n-1)} \\ &2x^n \frac{x^n}{1-x} - x^{2n} = x^{2(n-1)} \\ &2x^2 - x^2(1-x) = (1-x)\ , \end{aligned} \tag{13}$$

hence, x satisfies the cubic equation:

$$x^3 + x^2 + x - 1 = 0\ . \tag{14}$$

The only real solution to this equation is:

$$x = \lambda := \frac{1}{3}\left(-1 - \frac{2}{(17+3\sqrt{33})^{1/3}} + (17+3\sqrt{33})^{1/3}\right) \approx 0.54\ . \tag{15}$$

To ensure that $\sum_{k=1}^{\infty} p_k = 1$, we use the normalization condition:

$$\sum_{k=1}^{\infty} \alpha\lambda^{k-1} = 1\ . \tag{16}$$

Simplifying the geometric series, we find:

$$\alpha = 1 - \lambda \approx 0.46\ . \tag{17}$$

Thus, the Nash equilibrium strategy is:

$$p_n = (1-\lambda)\lambda^{n-1}, \tag{18}$$

where λ is the real solution of (14).

Uniqueness of solution

Up to now, we have provided a solution based on an ansatz, but we have not yet ruled out the possibility of other solutions. In this section, we will prove that the ansatz solution of equation (11) is, in fact, unique under the standard probability distribution hypotheses (i.e., non-negativity and normalization to 1).

The proof will proceed as follows:

1. **Recursive Structure**: The equilibrium probabilities satisfy a quadratic recurrence relation that determines each p_{n+1} based on the previous terms. This provides a clear framework to analyze the solution.

2. **Domain Definition**: The recursive relation imposes constraints to ensure that probabilities remain valid (i.e., non-negative) and that the square root term is well-defined. These constraints will bound p_1 within a specific interval.

3. **Monotonicity**: We will show, using strong induction, that the normalization as a function of the starting value $S(p_1) = \sum_{k=1}^{\infty} p_k$ has a strictly positive derivative within its valid domain. This, combined with the interval shaped domain for p_1 , ensures that this normalization function is a strictly increasing function of p_1 , and hence that the normalization condition $S(p_1) = 1$ can be satisfied by at most one value of p_1. Based on this, we will conclude that there exists a unique p_1 that leads to the solution $p_n = (1-\lambda)\lambda^{n-1}$, where λ is the parameter derived earlier.

So in the following we will proceed with the detailed proof.

Recursive Structure

First thing, we go back to equation (11) and we observe that any solution that constitutes a probability distribution must satisfy the normalization condition. Thus, equation (11) becomes,

$$p_{n+1}^2 - 2p_{n+1}\left(1 - \sum_{k=1}^{n} p_k\right) + p_n^2 = 0 \ . \tag{19}$$

We also remind that we are looking for solutions having $0 \le p_n \le 1$ for all n; this implies that the term $(1 - \sum_{k=1}^{n} p_k)$ must always be positive for such a solution to exist. Equation (19) defines a proper recursive relation, as knowing p_k for all the values of k up to n, allows us to compute p_{n+1}; we can do that by solving for p_{n+1} using the quadratic formula,

$$p_{n+1} = 1 - \sum_{k=1}^{n} p_k \pm \sqrt{\left(1 - \sum_{k=1}^{n} p_k\right)^2 - p_n^2} \ , \tag{20}$$

Only the minus sign allows to have $\left(\sum_{k=1}^{n+1} p_k\right) \le 1$ for all n, so we find

$$p_{n+1} = 1 - \sum_{k=1}^{n} p_k - \sqrt{\left(1 - \sum_{k=1}^{n} p_k\right)^2 - p_n^2} \ . \tag{21}$$

We observe that for any given value of p_1, the recurrence relation uniquely determines $p_2(p_1)$ (if it exists), which in turn uniquely defines $p_3(p_1)$ (if it exists), and so on. Consequently, for each starting value of p_1, there exists a unique (if valid) total sum $S(p_1) = \sum_{k=1}^{\infty} p_k(p_1)$ (for example, it's easy to see that we have $S(p_1 = 0) = 0$). Moreover, we have already established that when $p_1 = 1 - \lambda$, the solution satisfies $S(p_1 = 1 - \lambda) = 1$.
At this stage, however, it is possible that other values of p_1 might still satisfy $S(p_1) = 1$. We will show in the following that this function is strictly monotonically increasing in its domain of definition, thus implying the uniqueness of p_1 such that $S(p_1) = 1$. The first thing to do this is finding the domain of definition of the function $S(p_1)$.

Domain definition

The function $S(p_1)$ is well defined when all the $p_k(p_1)$ functions that appear in the sum are well defined[2] . We have that $p_1 \geq 0$ by request; the $p_k(p_1)$ are always well defined except where the arguments of the square roots in equations (21) become negative. The condition for positivity then (dropping for notational convenience the dependency on p_1) reads

$$\begin{aligned}
&\left(1-\sum_{k=1}^{n} p_k\right)^2 \geq p_n^2 \\
&\left(1-\sum_{k=1}^{n-1} p_k\right)^2 - 2p_n\left(1-\sum_{k=1}^{n-1} p_k\right) \geq 0 \\
&\left(1-\sum_{k=1}^{n-1} p_k\right) - 2p_n \geq 0
\end{aligned} \tag{22}$$

so we find

$$p_n \leq \frac{1}{2}\left(1-\sum_{k=1}^{n-1} p_k\right) \ . \tag{23}$$

Next, we are going to use formula (21) for the left hand side and, after some algebra, find an equation similar to equation (23) for p_{n-1}; it is therefore useful to define $\beta_1 = \frac{1}{2}$ and express equation (23) as

$$p_n \leq \beta_1\left(1-\sum_{k=1}^{n-1} p_k\right) \ . \tag{24}$$

[2]Also, we need the series to converge, but the choice of the minus sign in equation (21) trivially ensures it, as all the partial sums up to n are less than 1 .

Now, using equation (21), we find

$$
\begin{aligned}
&1-\sum_{k=1}^{n-1} p_k-\sqrt{\left(1-\sum_{k=1}^{n-1} p_k\right)^2-p_{n-1}^2} \leq \beta_1\left(1-\sum_{k=1}^{n-1} p_k\right) \\
&(1-\beta_1)\left(1-\sum_{k=1}^{n-1} p_k\right) \leq \sqrt{\left(1-\sum_{k=1}^{n-1} p_k\right)^2-p_{n-1}^2} \\
&(1-\beta_1)^2\left(1-\sum_{k=1}^{n-1} p_k\right)^2 \leq \left(1-\sum_{k=1}^{n-1} p_k\right)^2-p_{n-1}^2 \\
&p_{n-1}^2 \leq \left(1-(1-\beta_1)^2\right)\left(1-\sum_{k=1}^{n-1} p_k\right)^2 \\
&p_{n-1} \leq \sqrt{(1-(1-\beta_1)^2)}\left(1-\sum_{k=1}^{n-1} p_k\right) \\
&\left(1+\sqrt{(1-(1-\beta_1)^2)}\right) p_{n-1} \leq \sqrt{(1-(1-\beta_1)^2)}\left(1-\sum_{k=1}^{n-2} p_k\right) \\
&p_{n-1} \leq \left(\frac{\sqrt{(1-(1-\beta_1)^2)}}{1+\sqrt{(1-(1-\beta_1)^2)}}\right)\left(1-\sum_{k=1}^{n-2} p_k\right) \\
&p_{n-1} \leq \beta_2\left(1-\sum_{k=1}^{n-2} p_k\right)
\end{aligned}
\tag{25}
$$

If we iterate this computation m times, we find

$$
p_{n-m} \leq \beta_{m+1}\left(1-\sum_{k=1}^{n-m-1} p_k\right) \ , \tag{26}
$$

with β_m satisfying the recurrence relation

$$
\beta_{m+1} = \left(\frac{\sqrt{(1-(1-\beta_m)^2)}}{1+\sqrt{(1-(1-\beta_m)^2)}}\right) \ , \tag{27}
$$

where, we remind, $\beta_1 = 1/2$. Hence, we find that the domain of $p_n(p_1)$ is given by

$$
0 \leq p_1 \leq \beta_n \ . \tag{28}
$$

The sequence for β_n starting from $\beta_1 = 1/2$ is always decreasing, so the domain of $S(p_0)$ is given by

$$
0 \leq p_1 \leq \beta_\infty \ , \tag{29}
$$

where $\beta_\infty = \lim_{n\to\infty} \beta_n$. Then, β_∞ can be found by

$$
\begin{aligned}
\beta_\infty &= \left(\frac{\sqrt{(1-(1-\beta_\infty)^2)}}{1+\sqrt{(1-(1-\beta_\infty)^2)}} \right) \\
\frac{1}{\beta_\infty} &= 1 + \frac{1}{\sqrt{(1-(1-\beta_\infty)^2)}} \\
\left(\frac{1}{\beta_\infty} - 1 \right)^2 &= \frac{1}{1-(1-\beta_\infty)^2} \\
\frac{\beta_\infty^2}{(1-\beta_\infty)^2} &= 1-(1-\beta_\infty)^2 \\
\beta_\infty^2 &= (1-\beta_\infty)^2 \left(2\beta_\infty - \beta_\infty^2 \right)
\end{aligned}
\tag{30}
$$

$\beta_\infty = 0$ is a fixed point solution that is not the stable attractive one, so we can safely discard it. Hence, we are left with

$$
\beta_\infty = (1-\beta_\infty)^2 (2-\beta_\infty) \ . \tag{31}
$$

Finally, we introduce a change of variable, for reasons that will be clear in a moment, as $y = 1-\beta_\infty$. Therefore, the equation becomes

$$
\begin{aligned}
&1 - y = y^2(2-1+y) \\
&y^3 + y^2 + y - 1 = 0
\end{aligned}
\tag{32}
$$

where we recognize the same equation as (14), whose only real solution is $y = \lambda$, from which $\beta_\infty = 1-\lambda$, the same value as p_1 of our ansatz solution (18).

Monotonicity

We found that the function $S(p_1)$ is well defined for every p_1 such that $0 \le p_1 \le (1-\lambda)$, and we already know that if $p_1 = (1-\lambda)$, we have $S(p_1) = 1$. Now we will show that the function $S(p_1)$ is strictly monotonically increasing by computing its derivative and finding that it is strictly positive in its entire connected domain.

We can compute the derivative of $S(p_1)$ as

$$
S'(p_1) = \sum_{k=1}^{\infty} p_k'(p_1) \ . \tag{33}
$$

We will show, by strong induction, that $p_k'(p_1)$ is strictly positive for all k, thus proving that $S'(p_1)$ is strictly positive. The first step, $k = 1$, is trivial, as $p_1'(p_1) = 1$. Then, using equation (21), we find

$$
p_{n+1}'(p_1) = -\sum_{k=1}^{n} p_k'(p_1) - \frac{\left(1-\sum_{k=1}^n p_k(p_1)\right)\left(-\sum_{k=1}^n p_k'(p_1)\right) - p_n(p_1)p_n'(p_1)}{\sqrt{\left(1-\sum_{k=1}^n p_k(p_1)\right)^2 - p_n^2(p_1)}} \ . \tag{34}
$$

We can use again equation (21) to get rid of the square root in the denominator,

$$
p_{n+1}'(p_1) = -\sum_{k=1}^{n} p_k'(p_1) - \frac{\left(1-\sum_{k=1}^n p_k(p_1)\right)\left(-\sum_{k=1}^n p_k'(p_1)\right) - p_n(p_1)p_n'(p_1)}{1-\sum_{k=1}^{n+1} p_k(p_1)} \ . \tag{35}
$$

Rearranging the terms, we find

$$p'_{n+1}(p_1) = \left(\sum_{k=1}^{n} p'_k(p_1)\right)\left(-1 + \frac{1-\sum_{k=1}^{n} p_k(p_1)}{1-\sum_{k=1}^{n+1} p_k(p_1)}\right) + \frac{p_n(p_1)p'_n(p_1)}{1-\sum_{k=1}^{n+1} p_k(p_1)} \ . \tag{36}$$

The term $\left(\sum_{k=1}^{n} p'_k(p_1)\right)$ is positive, because of the inductive hypothesis. The term $\left(-1 + \frac{1-\sum_{k=1}^{n} p_k(p_1)}{1-\sum_{k=1}^{n+1} p_k(p_1)}\right)$ is positive because all the p_k are positive and we observed that for every n the partial sums give $\sum_k^n p_k(p_1) \leq 1$. Finally, the term $\left(\frac{p_n(p_1)p'_n(p_1)}{1-\sum_{k=1}^{n+1} p_k(p_1)}\right)$ is also positive due to the inductive hypothesis, the positivity of p_n and the the fact $\sum_k^{n+1} p_k(p_1) \leq 1$. Hence, we conclude that $p'_{n+1}(p_1) > 0$ and therefore $S'(p_1) > 0$.

To sum up, we have the following picture: $S(p_1)$ is an increasing function over the domain $0 \leq p_1 \leq 1-\lambda$, starting from $S(0) = 0$ and reaching at the extreme $S(1-\lambda) = 1$. Hence, the only solution compatible with positivity and the normalization condition is the one with $p_1 = 1-\lambda$, which takes the explicit form $p_n = (1-\lambda)\lambda^{n-1}$. We report in figure 1 a graphical representation of the function $S(p_1)$ obtained numerically.

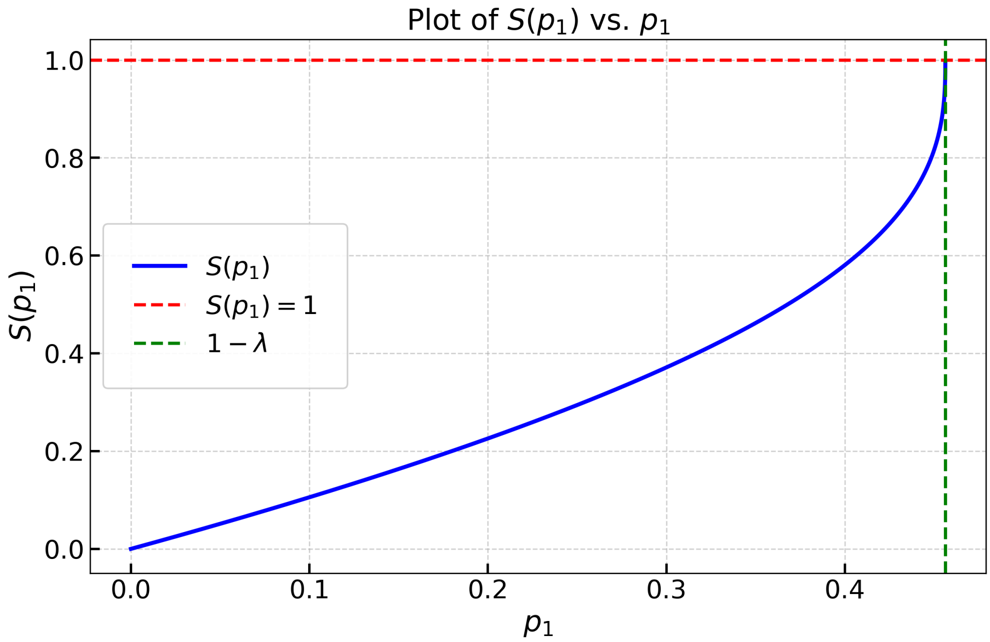

Figure 1: Plot of the function $S(p_1)$ versus p_1. The function is monotonic, and reaches $S(p_1) = 1$ for $p_1 = 1-\lambda$.

Conclusion

This document presented a solution to the final problem in the G-Research Sample Quant Exam, modeling it as a mixed-strategy game. We derived the Nash equilibrium equations, identified the unique exponential strategy parametrized by λ, and proved its uniqueness under positivity and normalization constraints.

The analysis highlights the symmetry among players and the exponential decay of probabilities, offering insights into strategic decision-making in combinatorial games. Future work could explore extensions with additional players or modified rules.

Appendix

The uniqueness of the solution is based on the restriction to symmetric solution for the three players. In this appendix we give a rigorous proof that solutions to the system of equations (8), (9) and (10) that satisfy the probability distribution hypotheses (i.e. positivity and normalization to 1) must be symmetric. To do so, we observe that these equations are bilinear and do not involve any quadratic term in the forms p_n^2, q_n^2 or r_n^2. Thus, we will suppose there is a (non necessarily symmetric) solution of the system where the p_n are equal to some non-zero p_n^* and we will use this linearity to show that p_n^* must be the solution also for q_n and r_n. As a preliminary step to do this, we will show the following lemma:

Lemma

Any solution of equations (8), (9) and (10) that satisfies the conditions for probability distributions cannot involve a sequence p_n, q_n or r_n that, for every $m > \overline{n}$, has $p_m = 0$, $q_m = 0$ or $r_m = 0$.

Proof

Assume, for contradiction, that p_n (r_n or q_n would be analogous) is such that for every $m > \overline{n}$ we have $p_m = 0$. Then, we consider N the last value of n for which $p_n \neq 0$, and we evaluate equations (9) and (10) in $n = N$. We immediately see that the left hand side of both is zero, and since $p_N \neq 0$, we must have that $r_N = 0$ from equation (9) and $q_N = 0$ from equation (10).

Now, if $N = 1$, we easily see that it must be that at least one of q_n and r_n is zero for all the values of n, contradicting the probability distribution hypothesis (either this, or we must have that q_2 and r_2 are both zero, and thus we can reiterate the reasoning.). If $N > 1$, by evaluating equation (8) in $n = N - 1$, we have that the left hand side becomes zero, and so we have that at least one between q_{N-1} and r_{N-1} must be zero. Suppose for simplicity that $q_{N-1} = 0$. We can therefore evaluate equation (10) (equation (9) in case $r_{N-1} = 0$) in $n = N - 1$. The right hand side is obviously zero; the first and third terms of the left hand side are also zero, because we found that $q_N = 0$; since we made the hypothesis that $p_N \neq 0$, we must therefore have that all $q_n = 0$ for all $n \geq N$, and actually $q_n = 0$ for all $n \geq N - 1$, because we already had that $q_{N-1} = 0$.

Now, if $N = 2$, we have that $q_n = 0$ for all N, contradicting the probability distribution hypothesis. If $N > 2$, we can then look at equation (9) and evaluate it at $n = N - 1$; since we already established that $r_N = 0$, and $p_N \neq 0$, wheter the left hand side is zero or not only depends on the term $\sum_{k=N}^{\infty} r_k$. The proof now splits in two:

- if $\sum_{k=N}^{\infty} r_k \neq 0$, we have that the left hand side of equation (9) at $n = N - 1$ is different from zero. Thus, we deduce that both p_{N-1} and r_{N-1} are different from zero. Hence, we can iterate the use of equation (9) for decreasing values of n until $n = 1$, and at every step we find that both p_n and r_n are greater than zero. Finally, we go back to equation (10) evaluated at $n = N - 2$; since we know that $q_n = 0$ for every $n \geq N - 1$, the left hand side is zero, and thus, since we found that $p_{N-2} \neq 0$, we find that $q_{N-2} = 0$. Iterating the reasoning using equation (10) for decreasing values of n, we find that $q_n = 0$ for all n, thus contradicting the probability distribution hypothesis.
- if $\sum_{k=N}^{\infty} r_k = 0$, we have that the left hand side of equation (9) at $n = N - 1$ is equal to zero and therefore we find that at least one between p_{N-1} and r_{N-1} is zero. At this point, we are left with the situation summarized in the table:

	1	$\cdots$	$N-2$	$N-1$	N	$N+1$	$N+2$	$\cdots$
p_n					X	0	0	$\cdots$
q_n				0	0	0	0	$\cdots$
r_n					0	0	0	$\cdots$

Where X represents non-zero values and empty squares are yet to be determined.

We will now show that at least one row has to contain only zeroes, thus contradicting the probability distribution hypothesis. We start from equation (9) evaluated at $n = N - 1$ and we immediately see that one between p_{N-1} and r_{N-1} (or both) must be zero, as the left hand side is zero. We have three possibilities:

- $r_{N-1} \neq 0$, and therefore $p_{N-1} = 0$. We can therefore look again at equation (9), evaluating it at $n = N - 2$; the left hand side contains a term $r_{N-1}p_N$ strictly greater than zero, so we find that the right hand side has to also be different from zero, and so both p_{N-2} and r_{N-2} are different from zero. We can iterate the use of equation (9) for decreasing values of n until $n = 1$, and at every step we find that both p_n and r_n are greater than zero. Thus, going back to equation (10) evaluated at $n = N - 2$, we find that the left hand side is zero, and so the right hand side must be zero as well. Since we established that $p_{N-2} \neq 0$, we must have that $q_{N-2} = 0$. Iterating the use of equation (10) for decreasing values of n until $n = 1$, we find $q_n = 0$ for all n, thus contradicting the probability distribution hypothesis.
- $p_{N-1} \neq 0$ and therefore $r_{N-1} = 0$. We can therefore evaluate equation (8) at $n = N - 2$, finding that at least one between q_{N-2} and r_{N-2} is zero. Thus, if $N-2 = 1$, we reach a row of all zeroes, contradicting the probability distribution hypothesis. If, instead, $N - 2 > 1$, we are left with a situation equivalent to the one represented in the previous table, but with a value of n lower by one, where the reasoning can be iterated until all the table is completed (or a contradiction is found). A table where q_{N-2} is the one equal to zero is shown below (r_{N-2} is equivalent).

	1	$\cdots$	$N-2$	$N-1$	N	$N+1$	$N+2$	$\cdots$
p_n				X	X	0	0	$\cdots$
q_n			0	0	0	0	0	$\cdots$
r_n				0	0	0	0	$\cdots$

- both p_{N-1} and r_{N-1} are equal to zero (and, we remind, so does q_{N-2}). Equations (8), (9) and (10) together, evaluated at $n = N - 2$ imply that at least two out of three among p_{N-2}, q_{N-2} and r_{N-2} are equal to zero. Since q_n and r_n are equivalent, we can assume for simplicity that $q_{N-2} = 0$, with p_{N-2} and r_{N-2} left to find. We observe, again, that if $N - 2 = 1$, we have $q_n = 0$ for all n, thus contradicting the probability distribution hypothesis. If $N - 2 > 1$, knowing that at least one among p_{N-2} and r_{N-2} is zero, we finally find that there are three sub-cases are possible:
  - both p_{N-2} and r_{N-2} are equal to zero. We are therefore in a situation analogous to the beginning of this step, with all p_{N-2}, q_{N-2} and r_{N-2} equal to zero. We can therefore reiterate the reasoning until the table is completed (or a contradiction is found).
  - $p_{N-2} \neq 0$ and $r_{N-2} = 0$. We immediately see, at this point, that this situation is equivalent to the one represented in the previously shown tables, and so we can iterate again the reasoning until the table is completed (or a contradiction is found).
  - $p_{N-2} = 0$ and $r_{N-2} \neq 0$. This case is represented in the table below for clarity, but is again somewhat similar to previously discussed possibilies.

	1	$\cdots$	$N-2$	$N-1$	N	$N+1$	$N+2$	$\cdots$
p_n			0	0	X	0	0	$\cdots$
q_n			0	0	0	0	0	$\cdots$
r_n			X	0	0	0	0	$\cdots$

  In this situation, we consider equation (9) evaluated in $n = N - 3$; the left hand side contains a term $r_{N-2}p_N$ strictly greater than zero, so we find that the right hand side has to also be different from zero, and so both p_{N-3} and r_{N-3} are different from zero. We can iterate the use of equation (9) for decreasing values of n until $n = 1$, and at every step we find that both p_n and r_n are greater than zero. Thus, going back to equation (10) evaluated at $n = N - 3$, we find that the left hand side is zero, and so the right hand side must be zero as well. Since we established that $p_{N-3} \neq 0$, we must have that $q_{N-3} = 0$. Iterating the use of equation (10) for decreasing values of n until $n = 1$, we find $q_n = 0$ for all n, thus contradicting the probability distribution hypothesis.

We see that the reasoning admits several splits due to multiple iterations, however, we have shown that all the possibilities will eventually lead to a contradiction. Hence, the statement of the lemma is proven.

Now that the lemma is proven, we can show that, indeed, solutions to the system of equations (8), (9) and (10) that satisfy the hypotheses of probability distribution must be symmetric.

As we observed earlier, these equations are bilinear and do not involve any quadratic term in the forms p_n^2, q_n^2 or r_n^2. Thus, let's suppose there is a (non necessarily symmetric) solution of the system where the p_n are equal to some p_n^* that satisfies the condition for a probability distribution. Equations (9) and (10) can therefore be seen as independent linear equation systems for the q_n and r_n variables. Using the normalization condition, they can be re-written as,

$$-p^*_{n+1}\sum_{k=1}^{n} r_k + r_{n+1}\sum_{k=n+1}^{\infty} p^*_k - p^*_{n+1}r_{n+1} - p^*_n r_n = -p^*_{n+1} \ , \tag{37}$$

$$-p^*_{n+1}\sum_{k=1}^{n} q_k + q_{n+1}\sum_{k=n+1}^{\infty} p^*_k - p^*_{n+1}q_{n+1} - p^*_n q_n = -p^*_{n+1} \ . \tag{38}$$

The left hand side of both the equations can be seen as the application of a linear (and identical, as it is uniquely defined by p^*_n) operator to the infinite-dimensional vectors r_n and q_n respectively. The right hand side, instead is a nun-null vector. Standard theorems of linear algebra, therefore, ensure that, for equation (37) and (38), either

- there is no solution;
- the space of all solutions constitutes an affine space (that is not a proper vector space including the null vector).

The first possibility is ruled out by the fact the we found at least one solution (the symmetric one of the main text) and we supposed that there is a solution with a non-null $p_n = p^*_n$. Therefore, we know that the solutions of equation (37) (equation (38) is analogous) constitutes an affine space of dimension $d \geq 0$. We now observe that equation (37) can be solved for r_{n+1} as

$$r_{n+1} = \frac{p^*_{n+1}\sum_{k=1}^{n} r_k + p^*_n r_n - p^*_{n+1}}{\left(\sum_{k=n+2}^{\infty} p^*_k\right)} \ , \tag{39}$$

Where we used the lemma to ensure the denominator is always different from zero. Equations (39) constitutes a set of recursive relations that allows us, given a certain value of p_1, to uniquely identify p_n for all $n > 1$. Hence, the affine space of the solutions of equation (37) has dimension $d = 1$. In geometric terms, this space a straight line in the infinite-dimensional space.
We now consider the normalization condition for the solution,

$$\sum_{k=1}^{\infty} r_k = 1 \ , \tag{40}$$

and we observe that it identifies an affine space that, in geometric terms, constitutes an hyperplane. Thus, if we aim for r_n that satisfies both equation (37) and the normalization condition, equation (40), we can have two possibilities:

- the hyperplane and the straight line have a single point intersection;
- the straight line is parallel and belongs to the hyperplane, and thus the intersection is the straight line itself (the possibility of no intersection parallel line is ruled out because we know at least one normalizable solution exists).

Only in the first case, there is a unique normalizable solution. We can however show that the line does not belong to the hyperplane finding one point of it that doesn't belong to the plane. To find such a point, is enough to consider equation (39) and take any $r_1 > 1$. Now, it is clear that the recursive relation will only give us positive values for r_n for all n, as p_n is a probability distribution. Thus, we find that $\sum_k r_k > 1$ and we conclude that, among all the solutions of equation (37), there is only one that is normalizable to 1.

The fact that equation (37), given a certain p_n^*, has a unique normalizable solution means that equation (38) has also a unique normalizable solution, and it must be the same (although,for now they might be different from p_n^*). Let's call this solution $r_n = q_n = q_n^*$.
Making again the whole reasoning considering, this time, q_n^* as the given part of the solution for the equation system (8), (9) and (10), we find that the unique solutions for the linear system gives $p_n = r_n$. But since we know that q_n^* is the solution for the q_n and the r_n that we have when p_n^* is the solution for the p_n, we conclude that all q_n, r_n and p_n have the same solution $q_n^* = p_n^*$, that is therefore the one found in the main text in equation (18) .

References

[1] G-Research, *Sample Quant Exam*, 2017. Available at: `https://www.gresearch.com/wp-content/uploads/2019/12/Sample-Quant-Exa.pdf`