

with K categories $\{1, 2, \dots, K\}$ with $P(\varepsilon_1 = j) = p_j, j = 1, \dots, K$, with $\sum_{j=1}^K p_j = 1$. Define the *dependency coefficient* $\delta \in [0, 1]$. Korzeniowski [1] and Traylor [2] defined the following quantities:

$$\begin{aligned} p_j^+ &:= p_j + \delta \left(\sum_{i \neq j} p_i \right) = p_j + \delta(1 - p_j) \\ p_j^- &:= p_j - \delta p_j \end{aligned} \quad \text{for } j = 1, 2, \dots, K \quad (1)$$

where we have the following identities.

$$p_j^+ + \sum_{i \neq j} p_i^- = 1 \text{ for } j = 1, 2, \dots, K \quad (2)$$

Under first-kind dependence,

$$\begin{aligned} P(\varepsilon_i = j | \varepsilon_1 = j) &= p_j^+ \\ P(\varepsilon_i = j | \varepsilon_1 \neq j) &= p_j^- \end{aligned} \quad \text{for } j = 1, 2, \dots, K; \quad i \geq 2 \quad (3)$$

Figure 1 gives an illustration of standard first-kind dependence for a sequence of categorical random variables with 3 categories. The binary tree shows the allocation of probability mass at each new variable in the sequence given the previous outcomes. Note that standard first-kind dependence weights the probability of subsequent outcomes *in favor* of the outcome of the first random variable, and *away* from the others. As an explicit example, the probability that $\varepsilon_3 = 2$ is p_2^+ if $\varepsilon_1 = 2$, but p_2^- if $\varepsilon_1 = 1$ or 3.

Traylor [2] proved that $P(\varepsilon_i = j) = p_j$ for all $i = 1, 2, \dots$ and for all $j = 1, 2, \dots, K$. That is, despite the dependence of subsequent categorical variables in the sequence on the first, all variables remain identically distributed. Traylor and Hathcock [3] also extended this notion to other types of *vertical dependence* by creating a class of *vertical dependency generators* \mathcal{C} that generate other dependency structures in addition to first-kind dependence, such as sequential dependence. Any dependency structure generated by a function in \mathcal{C} , defined by $\mathcal{C}_\delta = \{\alpha : \mathbb{N}_{\geq 2} \rightarrow \mathbb{N}\}$ such that

- $\alpha(n) < n$
- $\forall n \exists j \in \{1, \dots, n-1\} : \alpha(n) = j$

The generating function for first-kind dependence is $\alpha(n) \equiv 1$, showing that every categorical random variable, starting with the second, assigns conditional probabilities based on the outcome of the first.

Regardless of vertical dependency structure, we have thus far assumed that we weight in favor of the outcome of $\alpha(n)$. That is, if $\varepsilon_{\alpha(n)} = j$, then $P(\varepsilon_n = j | \varepsilon_{\alpha(n)} = j) = p_j^+$, and $P(\varepsilon_n = j | \varepsilon_{\alpha(n)} \neq j) = p_j^-$. Under these conditions, any dependent categorical sequence generated by a function $\alpha \in \mathcal{C}$ is identically distributed but dependent, as proven in [3].

This paper examines permuted weights and derives the conditions under which first-kind dependent categorical random variables remain identically distributed under permuted weights. We define permuted weighting in the next section.

1. Permuted Weighting

Suppose we take a permutation σ of $\{1, 2, \dots, K\}$, so that $\sigma : i \mapsto \sigma_i, i = 1, \dots, K$. Now suppose that we take a first-kind dependent sequence of categorical random variables. Now, we define the permuted weights by modifying the conditional probabilities. For $i \geq 2$

$$\begin{aligned} P(\varepsilon_i = j | \varepsilon_1 = \sigma_j^{-1}) &= p_j^+ \\ P(\varepsilon_i = j | \varepsilon_1 \neq \sigma_j^{-1}) &= p_j^- \end{aligned} \quad \text{for } j = 1, 2, \dots, K \quad (4)$$

The *vertical dependency* - first kind dependency in this case- is unchanged, in that $\alpha(n) \equiv 1$, but *how* those specific probability masses are assigned has been permuted.

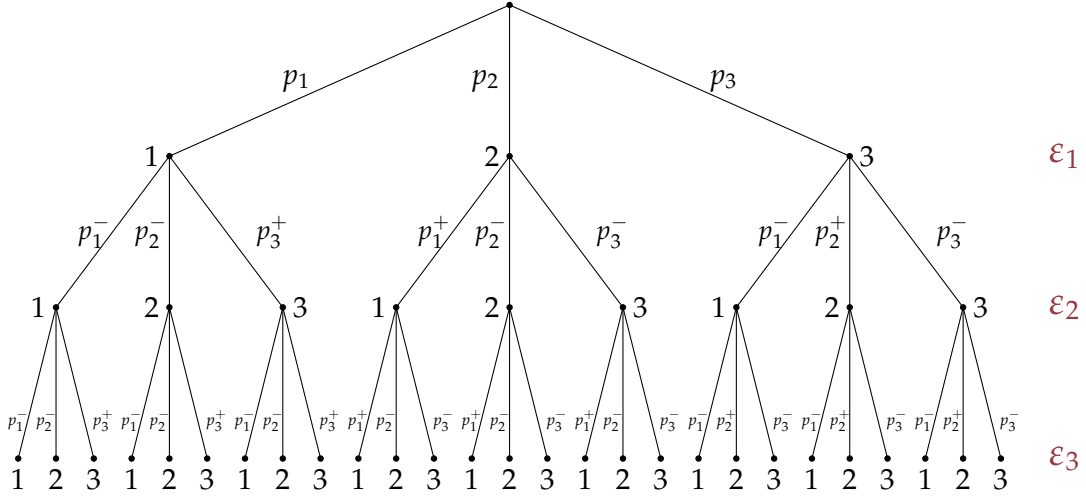


Figure 2. Permuted First Kind Dependence for Categorical Random Variables with $K = 3$, $\sigma = (132)$

Example 1. For example, suppose $K = 3$, and let $\sigma = (132)$. Then for $i \geq 2$,

$$\begin{aligned}
 P(\varepsilon_i = 1 | \varepsilon_1 = 1) &= p_1^- & P(\varepsilon_i = 1 | \varepsilon_1 = 2) &= p_1^+ & P(\varepsilon_i = 1 | \varepsilon_1 = 3) &= p_1^- \\
 P(\varepsilon_i = 2 | \varepsilon_1 = 1) &= p_2^- & P(\varepsilon_i = 2 | \varepsilon_1 = 2) &= p_2^- & P(\varepsilon_i = 2 | \varepsilon_1 = 3) &= p_2^+ \\
 P(\varepsilon_i = 3 | \varepsilon_1 = 1) &= p_3^+ & P(\varepsilon_i = 3 | \varepsilon_1 = 2) &= p_3^- & P(\varepsilon_i = 3 | \varepsilon_1 = 3) &= p_3^-
 \end{aligned} \tag{5}$$

The tree for first-kind dependence with weights according to $\sigma = (132)$ is given in Figure 2.

The following lemma gives the unconditional probability distribution of any categorical random variable $\varepsilon_i, i \geq 2$ in a permuted first-kind dependent sequence. Note that $P(\varepsilon_1 = j) = p_j$ by definition.

Lemma 1. Let $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ be a first-kind dependent sequence of categorical random variables with permuted weights given by σ . Then

$$P(\varepsilon_n = i) = p_{\sigma^{-1}(i)} p_i^+ + p_i^- \sum_{j \neq \sigma^{-1}(i)} p_j; \quad n > 1, \quad i = 1, 2, \dots, K$$

Proof. Let Ω_n denote the set of possible first-kind dependent categorical sequences of length n . WLOG, let $i = 1$. Denote the subsets of Ω_n where $\varepsilon_n = 1$ by $\Omega_n^{(1)}$. That is,

$$\Omega_n^{(1)} = \{\varepsilon \in \Omega_n : \varepsilon_n = 1\} \subset \Omega_n$$

Then we may decompose $\Omega_n^{(1)}$ as $\Omega_n^{(1)} = \Omega_{n-1} \times \{1\}$. Similarly, let $\Theta_n^{(i)} = \{\varepsilon \in \Omega_n : \varepsilon_1 = i\}$ and $\Omega_{n,-1}$ denote the first-kind dependent categorical sequences of length n with ε_1 deleted. Then we may express $\Theta_{n-1}^{(i)}$ as $\Theta_{n-1}^{(i)} = \{i\} \times \Omega_{n-1,-1}$. Therefore,

$$\Omega_n^{(1)} = \bigcup_{i=1}^K [\{i\} \times \Omega_{n-1,-1} \times \{1\}]$$

Thus,

$$\begin{aligned}
 P(\varepsilon_n = 1) &= P(\Omega_n^{(1)}) \\
 &= P\left(\bigcup_{i=1}^K [\{i\} \times \Omega_{n-1,-1} \times \{1\}]\right) \\
 &= p_{\sigma^{-1}(1)} P(\Omega_{n-1,-1}) p_1^+ + \sum_{i \neq \sigma^{-1}(1)} p_i P(\Omega_{n-1,1}) p_1^- \\
 &= p_{\sigma^{-1}(1)} p_1^+ + \sum_{i \neq \sigma^{-1}(1)} p_i p_1^-
 \end{aligned}$$

Note $P(\Omega_{n-1,1}) = 1$. A similar argument follows for $j = 2, \dots, K$. □

This lemma gives the unconditional probability distribution for the subsequent categorical random variables after ε_1 under permuted first-kind dependence. If σ is the identity permutation, then $\sigma^{-1}(i) = i$, and

$$P(\varepsilon_n = i) = p_i p_i^+ + p_i^- \sum_{j \neq i} p_j = p_i$$

and we recover standard first kind dependence from [2]. Now the entire sequence is identically distributed. In general, this may not always be true. The following theorem gives necessary and sufficient conditions on the initial probability distribution to guarantee an identically distributed permuted sequence of first-kind dependent categorical random variables.

Theorem 1. *Suppose $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ is a sequence of first-kind dependent categorical random variables with K categories. Let σ be a permutation of $\{1, 2, \dots, K\}$ that permutes the conditional probabilities of $\varepsilon_j, j \geq 2$ according to the outcome of ε_1 . That is, suppose*

$$P(\varepsilon_j = k | \varepsilon_1 = \sigma^{-1}(k)) = p_k^+ \text{ and } P(\varepsilon_j = k | \varepsilon_1 \neq \sigma^{-1}(k)) = p_k^-, \quad k = 1, 2, \dots, K$$

Then all variables ε_i are identically distributed with $P(\varepsilon_i = k) = p_k, k = 1, 2, \dots, K$ if and only if $p_i = p_{\sigma^{-1}(i)}$.

Proof. By Lemma 1,

$$\begin{aligned}
 P(\varepsilon_n = k) &= p_{\sigma^{-1}(k)} p_k^+ + p_k^- \sum_{i \neq \sigma^{-1}(k)} p_i \\
 &= (p_k + \delta \sum_{l \neq k} p_l) p_{\sigma^{-1}(k)} + (p_k - \delta p_k) \sum_{i \neq \sigma^{-1}(k)} p_i
 \end{aligned} \tag{6}$$

If $\sigma^{-1}(k) = k$, then we are back to Lemma 2 in [2] and finished. Therefore, suppose $\sigma^{-1}(k) \neq k$. Then

$$\begin{aligned}
 P(\varepsilon_n = k) &= (p_k + \delta \sum_{l \neq k} p_l) p_{\sigma^{-1}(k)} + (p_k - \delta p_k) \sum_{i \neq \sigma^{-1}(k)} p_i \\
 &= p_k p_{\sigma^{-1}(k)} + p_k \sum_{i \neq \sigma^{-1}(k)} p_i + \delta p_{\sigma^{-1}(k)} \left(\sum_{l \neq k} p_l \right) - \delta p_k \sum_{i \neq \sigma^{-1}(k)} p_i \\
 &= p_k + \delta p_{\sigma^{-1}(k)}^2 + \delta p_{\sigma^{-1}(k)} \sum_{\substack{l \neq k \\ l \neq \sigma^{-1}(k)}} p_l - \delta p_k^2 - \delta p_k \sum_{\substack{l \neq k \\ l \neq \sigma^{-1}(k)}} p_l \\
 &= p_k + \delta \left(p_{\sigma^{-1}(k)}^2 - p_k^2 \right) + \delta (p_{\sigma^{-1}(k)} - p_k) \sum_{\substack{l \neq k \\ l \neq \sigma^{-1}(k)}} p_l \\
 &= p_k + \delta (p_{\sigma^{-1}(k)} - p_k) \sum_{l=1}^K p_l \\
 &= p_k + \delta (p_{\sigma^{-1}(k)} - p_k)
 \end{aligned}$$

Thus $P(\varepsilon_n = k) = p_k$ if and only if $p_{\sigma^{-1}(k)} = p_k$ for all $k = 1, 2, \dots, K$. \square

Returning to the example above, where $K = 3$ and $\sigma = (132)$, we have by Theorem 1 that the entire sequence is identically distributed if and only if $p_1 = p_{\sigma^{-1}(1)} = p_2$, $p_2 = p_3$, and $p_3 = p_1$. In this case, we may conclude that only of $p_1 = p_2 = p_3 = p$ will the sequence remain identically distributed under first-kind dependence.

Example 2. Suppose now that $K = 4$, and $\sigma = (12)(34)$. Then by Theorem 1, a sequence ε of these categorical random variables under first-kind dependence will be identically distributed if and only if $p_1 = p_2$ and $p_3 = p_4$. Thus, in this case, $p_1 = p_2 = p$, and $p_3 = p_4 = p'$, but p and p' need not necessarily be equal to ensure the sequence is identically distributed.

This example illustrates a corollary to Theorem 1. A permuted FK-dependent sequence of categorical random variables is identically distributed if and only if the probabilities within each disjoint cycle are equal.

Corollary 1. Let $\sigma = \tau_1 \tau_2 \cdots \tau_m$ be a permutation decomposed into a product of disjoint cycles τ_j . Suppose ε is a first-kind dependent sequence of categorical random variables with weights permuted according to σ . Then the sequence is identically distributed if and only if the probabilities within each disjoint cycle are equal.

Proof. From Theorem 1, to ensure an identically distributed sequence, we must have that $p_k = p_{\sigma^{-1}(k)}$ for all $k = 1, 2, \dots, K$. Let τ_j be the orbit of some $j \in \{1, 2, \dots, K\}$, and suppose the order of j is $\gamma = |\tau_j|$. Then $\tau_j = (j \ \sigma(j) \ \sigma^2(j) \ \dots \ \sigma^{\gamma-1}(j))$. By Theorem 1, the sequence is identically distributed if and only if $p_j = p_{\sigma(j)} = \dots = p_{\sigma^{\gamma-1}(j)}$ within each cycle. \square

2. Remarks and Conclusion

The notion of permuted weights extends the possibilities for dependence among a sequence of categorical random variables beyond even the class of vertical dependency structures. This paper introduced the notion of *horizontal dependence* via permuting the modification of the subsequent outcome probabilities, giving the possibility of two different sequences with the same vertical dependency structure, but different horizontal dependency. Permuting the conditional probabilities away from the identity permutation requires the addition of restrictions on the initial probability distribution to retain the desired identically distributed nature of the sequence. Theorem 1 and Corollary 1 show that the necessary and sufficient

condition to guarantee an identically distributed sequence of first-kind dependent categorical random variables is for the probabilities within each cycle to be equal.

The identity permutation $i = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = (1)(2) \cdots (n)$ when decomposed into a product of disjoint cycles. Thus, we can see that this is a special case of Theorem 1 and Corollary 1. First-kind dependent sequences of categorical random variables under the identity permutation are always identically distributed, as shown in [2]. In fact, vertical dependency structures generated by $\alpha \in \mathcal{C}$ from [3] are also always identically distributed under the identity permutation. At the other extreme, suppose ε is a first-kind dependent sequence with permuted weights according to σ which is a 1-cycle. Then the only way this sequence can be identically distributed is if $p_1 = \cdots = p_K$. Future research will examine necessary and sufficient conditions on horizontal dependence to retain the identically distributed property of sequences of categorical random variables in other dependency structures.

References

- [1] Andrzej Korzeniowski. On correlated random graphs. *Journal of Probability and Statistical Science*, pages 43–58, 2013.
- [2] Rachel Traylor. A generalized multinomial distribution from dependent categorical random variables. *Academic Advances of the CTO*, 2017.
- [3] Rachel Traylor and Jason Hathcock. Vertical dependency in sequences of categorical random variables. *AACTO*, 2017.