Order Statistics from Independent Non-Identical Exponentiated and Proportional Hazard Rate Random Variables

José-Antonio Espín-Sánchez
*Yale University*, jose-antonio.espin-sanchez@yale.edu

Tianhao Wu
*Yale University*, tianhao.wu@yale.edu

Follow this and additional works at: [https://elischolar.library.yale.edu/cowles-discussion-paper-series](https://elischolar.library.yale.edu/cowles-discussion-paper-series)

Part of the Economics Commons

**Recommended Citation**
[https://elischolar.library.yale.edu/cowles-discussion-paper-series/2707](https://elischolar.library.yale.edu/cowles-discussion-paper-series/2707)

This Discussion Paper is brought to you for free and open access by the Cowles Foundation at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Cowles Foundation Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elisorhale@yale.edu.
ORDER STATISTICS FROM INDEPENDENT NON-IDENTICAL EXPONENTIATED AND PROPORTIONAL HAZARD RATE RANDOM VARIABLES*

By

Jose-Antonio Espin-Sanchez and Tianhao Wu

September 2022

COWLES FOUNDATION DISCUSSION PAPER NO. 2346

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/
Order Statistics from Independent Non-Identical Exponentiated and Proportional Hazard Rate Random Variables

José-Antonio Espín-Sánchez† Tianhao Wu‡

September 2, 2022

Abstract

We study order statistics (OS) from independent non-identically distributed (INID) samples for two large classes of statistical distributions: Exponentiated Distributions (ED) and Proportional Hazard Rate Models (PHRM). We show that for the analytical solution for the CDF (PDF) of OSs in ED and PHRM: i) each OS’s CDF (PDF) depends on all shape parameters; ii) the CDF (PDF) of each OS is a weighted average of CDF (PDF) within the same family and with shape parameters equal to a partial sum of the original shape parameters; and iii) the weights are integers and sum up to 1. These properties allows for a clear analytical solution and allows a simple parameter estimation in these classes of distributions.

JEL Codes: C4, C57, F11, L1
Keywords: Order Statistics, Maximum, Gumbel

†We appreciate helpful comments from Timothy Armstrong, Steve Berry, Jorge Catepillan, Charles Hodgson, Nicholas Ryan, Mykhaylo Shkolnikov, Santiago Truffa, and Chris Vickers.
‡Department of Economics, Yale University, jose-antonio.espin-sanchez@yale.edu.
§Department of Economics, Yale University, tianhao.wu@yale.edu.
1 Introduction

Order statistics (OS) are widely used in statistics and economics. The main applications are auction models, discrete-choice demand estimation, macro/trade models and survival analysis. First, in a second price auction, each bidder submits a bid equal to their valuation. The winning price is a random variable that follows the $(n - 1)^{th}$ OS of the underlying distribution of valuations (Krishna, 2002; Espín-Sánchez et al., forthcoming). Second, in discrete-choice demand estimation models, the consumer chooses the good that provides the highest utility. This utility is a random variable that follows the $n^{th}$ OS (maximum) among the utilities provided by each good (Berry et al., 1995; Grieco et al., 2022). Third, in Ricardian trade models firms have different efficiency levels and only the firm with the highest efficiency (maximum) would produce (Eaton and Kortum, 2002). Fourth, in survival analysis with competing risks, the time until exit follows the first order statistic (minimum) among the exit times for each risk type (Cox and Oakes, 1984; Dickman et al., 2004).

In this article, we compute the analytic formulation of the cumulative probability function (CDF) and probability density function (PDF) of all OS for independent non-identically distributed (INID) samples generated by two widely used classes of distributions: Exponentiated Distributions (ED) with CDF $F(x, \theta) = G(x)^\theta$ and Proportional Hazard Rate Models (PHRM) with CDF $F(x, \theta) = 1 - [1 - G(x)]^\theta$, where $G(x)$ is any CDF. We show that the CDF (PDF) of any order statistic within these classes can be written as a weighted average of CDF (PDF) of the same class, where each CFD (PDF) have parameters that are a partial sum of the parameters from the original distributions. Moreover, the weights are integers that sum up to 1. This last property is unique to OS from the classes studied here. This makes parameter estimation by Maximum Likelihood simple and fast.

The Gumbel distribution is widely used in economics for discrete-choice demand estimation due to its analytic properties (Berry et al., 1995). The maximum of a set of random variables that follow a Gumbel distribution, each with the same scale parameter but different location parameter, is also a random variable that follows a Gumbel distribution with the same scale parameter, and a location parameter that is a function of the original location parameters. Having an analytical solution for the distribution of the maximum simplifies the analysis of

\footnote{Bernard et al. (2003) assume a form of imperfect competition where the most efficient firm produces at a price dependent on the second most efficient firm.}
discrete choice models. The Gumbel distribution is a particular case of an ED (Froeb et al., 2001). Our results extend the analytic properties of the maximum of a set of variables distributed Gumbel to variables distributed according to any ED. Moreover, we provide analytic solutions for all OS, not only the maximum, for the Gumbel distribution and any other ED. Whereas the distribution of the maximum is interesting in discrete-choice models where consumers can buy at most one unit, the distribution of the $r^{th}$ order statistic is useful when consumers can buy up to $(n - r)^{th}$ units.

OS from independent and identically distributed (IID) samples are well studied. The IID assumption is natural in settings where the data is generated by an homogeneous population, or the econometrician does not have additional information. Shakil and Ahsanullah (2012) computes the moments of OS for ED.\footnote{In the statistics literature, particular cases of ED are used and scholars compute the moments of their OS: generalized ED (Raqad and Ahsanullah, 2001); Exponentiated Gamma (Shawky et al., 2009).} Our results generalize existing results in the literature by providing an analytical solution to the OS’s CDF (and PDF), not only on IID samples but also on INID samples as explained below.

There are settings, however, where the statistician knows other characteristics of the population and can classify observations as belonging to distinct groups. For example, in an auction for timber, the statistician may know which bids belong to big firms (millers) vs small firms (loggers) (Athey et al., 2011; Roberts and Sweeting, 2013). In such settings, knowing the properties of OS from non-identical distributions would help in the estimation of the parameters. OS from INID, however, have received less scholarly attention.\footnote{The statistics literature that studies OS typically focuses on relations of product moments and stochastic orderings: Balakrishnan (1994) on exponential random variables, Childs et al. (2001) on right-truncated Lomax random variables, and Balakrishnan and Balasubramanian (1995) on power function random variables.}

2 Setting

In this section, we present the relevant definitions and auxiliary results in the literature.

**Definition (ED).** The CDF $F(x, \alpha, \theta_i)$ of a random variable $X_i$ which follows an
Exponentiated Distribution (ED) is defined by

\[ F(x, \alpha, \theta_i) = G(x, \alpha)^{\theta_i}, \theta_i > 0, \] (1)

where \( G(x, \alpha) \) is the source CDF, \( \theta_i \) is a positive shape parameter and \( \alpha \) is the source parameter vector. Denote the corresponding PDF of \( G(x, \alpha) \) as \( g(x, \alpha) \), we can write the PDF of an ED random variable

\[ f(x, \alpha, \theta_i) = \theta_i g(x, \alpha) G(x, \alpha)^{\theta_i-1}, \theta_i > 0. \] (2)

The ED, or power-related distribution, is sometimes called the Proportional Reversed Hazard Rate Model (PRHRM) (Gupta and Gupta, 2007; Chen and Xu, 2015). The reversed hazard rate of \( F \) is \( \frac{d}{dx} \log(F) = f/F = \theta g/G = \theta \frac{d}{dx} \log(G) \), which implies that the reversed hazard rate of \( F \) is proportional to that of \( G \), and to that of any other distribution \( F \) with the same source distribution \( G \).

Definition (PHRM). The CDF \( F(x, \alpha, \theta_i) \) of a random variable \( X_i \) which follows a Proportional Hazard Rate Model (PHRM) is defined by

\[ 1 - F(x, \alpha, \theta_i) = [1 - G(x, \alpha)]^{\theta_i}, \theta_i > 0, \] (3)

where \( G(x, \alpha) \) is the source CDF, \( \theta_i \) is a positive shape parameter and \( \alpha \) is the source parameter vector. Denote the corresponding PDF of \( G(x, \alpha) \) as \( g(x, \alpha) \), we can write the PDF of a PHRM random variable

\[ f(x, \alpha, \theta_i) = \theta_i g(x, \alpha)[1 - G(x, \alpha)]^{\theta_i-1}, \theta_i > 0. \] (4)

The hazard rate of a PHRM is \( \frac{f}{1-F} = \frac{\theta g[1-G]^{\theta-1}}{[1-G]^{\theta}} = \frac{\theta g}{1-G} \) which implies that the hazard rate of \( F \) is proportional to that of \( G \), and to that of any other distribution \( F \) with the same source distribution \( G \).

We now reproduce equation 5.2.1 from David and Nagaraja (2003), which shows the general formulation for OS from INID samples.\(^4\) We use this formulation for our results, and later we will show how the formulations from the ED and PHRM families are simpler. Suppose \( X_1, X_2, ..., X_n \) are independent random variables. Denote the CDF and PDF of \( X_i \) as \( F_i(x) \) and \( f_i(x) \) respectively. Let

\[^4\text{Also equation 1.1 in Balakrishnan and Balasubramanian (1995) and equations 2.1 and 2.2 in Jamjoom and Al-Saify (2012).}\]
\( X_{(1)}, X_{(2)}, ..., X_{(n)} \) denote the order statistics with an increasing order. Then, the cumulative distribution function of \( X_{(r)} \) \( (r \in \{1, 2, 3, ..., n\} \) can be written as

\[
F_r(x) = \sum_{i=r}^{n} \sum_{S_i} \prod_{k=1}^{i} F_{j_k} \prod_{k=i+1}^{n} [1 - F_{j_k}],
\]

(5)

where \( S_i \) extends all the permutations of \((j_1, j_2, ..., j_n)\) of \(1, 2, ..., n\) such that \(j_1 < j_2 < ... < j_i\) and \(j_{i+1} < j_{i+2} < ... < j_n\). The probability density function of \( X_{(r)} \) \( (r \in \{1, 2, 3, ..., n\} \) can be written as

\[
f_r(x) = \frac{1}{(r-1)! (n-r)!} \sum_S F_{j_1}...F_{j_{r-1}}f_{j_r} (1 - F_{j_{r+1}})...(1 - F_{j_n}),
\]

(6)

where \( S \) extends all the \( n! \) permutations of \((j_1, j_2, ..., j_n)\) of \(1, 2, ..., n\).

Equation 6 shows the PDF of the \( r^{th} \) order statistic as a weighted average of the PDF of each permutation \( j_r \). Notice that this weighted average depends both on the combinatorial probability outside the summation, and in the continuous probability inside the summation, i.e., the products of CDFs below \( r \) and the products of survival functions \( H(x) \) \( (H(x) = 1 - F(x)) \) above \( r \). Whereas the first term is a ratio of natural numbers which can be easily computed, the term inside the summation could depend on the CDFs of all random variables. In other words, in general, the weights assigned to each PDF depend (non-linearly) on the parameters to be estimated. This makes the use of the general formula very cumbersome for estimation.

3 Order Statistics

3.1 Exponentiated Distributions

We now show our results related to Exponentiated Distributions (ED) and we discuss their properties. Examples of ED distributions include Generalized Pareto type II (Lomax), Gumbel and Fréchet distributions.\(^5\) Proposition 1 shows our main result.

**Proposition 1 (ED).** Suppose \( X_1, X_2, ..., X_n \) are INID ED random variables with

---

\(^5\)The usual formulation for the Gumbel is \( G(x, \sigma, \mu) = \exp (-e^{-(x-\mu)/\sigma}) \), but we can also define \( \theta_i = e^{\mu_i}/\sigma \) and get \( F(x, \sigma, \theta_i) = [G(x, \sigma, 0)]^{\theta_i} = \left[e^{-e^{-\sigma(x-\mu_i)}}\right]^{\theta_i} = e^{-e^{-\sigma(x-\mu_i)}} = G(x, \sigma, \mu_i) \).
source CDF $G(x, \alpha)$ and shape parameter $\theta_1, \theta_2, ..., \theta_n$. Denote the CDF and PDF of $X_i$ as $F(x, \alpha, \theta_i)$ and $f(x, \alpha, \theta_i)$, which are defined in (1) and (2). Let $X_{(r)} (r \in \{1, 2, 3, ..., n\})$ denote the associated order statistics in ascending order. The CDF of $X_{(r)}$ can be written as

$$F_r(x) = \sum_{i=0}^{n-r} p_i(n, r) \sum_{S_{r+i}} F\left(x, \alpha, \sum_{k=1}^{r+i} \theta_{j_k}\right).$$ (7)

The PDF of $X_{(r)}$ can be written as

$$f_r(x) = \sum_{i=0}^{n-r} p_i(n, r) \sum_{S_{r+i}} f\left(x, \alpha, \sum_{k=1}^{r+i} \theta_{j_k}\right),$$ (8)

where $p_i(n, r) = (-1)^i C_r^{i+C} C_r^n = \frac{n!}{r!(n-r)!}$ and $S_{r+i}$ extends all combinatorics $(j_1, j_2, ..., j_{r+i})$ of $r+i$ numbers out of $1, 2, ..., n$.

**Proof.** See Appendix A.1. □

Equations 7 and 8 are much simpler than their counterparts in the general case, i.e., equations 5 and 6 respectively. In equation 8, the weights do not depend on the distributions or the parameters. The weights are just integer numbers. This means that we can use Maximum Likelihood estimation very efficiently. The density of an OS from ED is a weighed average of the CDF (PDF) from the same family, where the shape parameter in each CDF (PDF) is a partial sum of the parameters in the original distribution. The following corollary shows that not only are the weights integers and do not depend on the distributions or their parameters, but they sum up to 1.

**Corollary 1 (Weights).** The weights in the PDF (8) of the OS $X_{(r)}$ from $n$ independent non-identical ED random variables with shape parameters $\theta_1, \theta_2, ..., \theta_n$, do not depend on $\theta_1, \theta_2, ..., \theta_n$ and sum up to 1.

**Proof.** See Appendix A.2. □

We understand the notation to be cumbersome. Thus, we now show an example with $n = 3$ of a general ED to reinforce the intuition behind this elegant result.

**Example 1 (ED $n = 3$).** Suppose we have three ED random variables with shape parameters $\theta_1$, $\theta_2$ and $\theta_3$, i.e., $F(x, \theta_i) = [G(x)]^{\theta_i}$. We can write the PDF of the
minimum \( (r = 1) \), median \( (r = 2) \) and maximum \( (r = 3) \) as

\[
f_1(x) = \sum_{i=1}^{3} f(x, \theta_i) - f(x, \theta_1 + \theta_2) - f(x, \theta_1 + \theta_3) - f(x, \theta_2 + \theta_3) + f \left( x, \sum_{i=1}^{3} \theta_i \right),
\]

\[
f_2(x) = f(x, \theta_1 + \theta_2) + f(x, \theta_1 + \theta_3) + f(x, \theta_2 + \theta_3) - 2f \left( x, \sum_{i=1}^{3} \theta_i \right),
\]

\[
f_3(x) = f \left( x, \sum_{i=1}^{3} \theta_i \right).
\]

The PDF of the maximum \( (f_3(x)) \) in this example has a very simple solution: it is equal to the PDF of the same family, but the shape parameter is the sum of the original three shape parameters. This result may remind the reader of a similar widely known result: the maximum of several Gumbel variables is also a Gumbel. The Gumbel Distribution is one widely used example of ED, i.e., any Gumbel distribution can be written as an ED where the source distribution is a Gumbel \( G \) with location parameter \( \mu = 0 \).\(^6\) What the example here shows, and we prove below in Lemma 1, is that the common known result for the Extreme Types distributions, applies more generally to any ED. In order to get the nice analytical solution from the maximum of several Gumbels, the original distributions may have different location parameters \( \mu_i \) but they must have the same scale parameter \( \sigma \). The same caveat also applies here more generally. The source distribution \( G(x) \) could be a function of a vector of parameters \( \alpha \), i.e., \( G(x, \alpha) \), but this vector of parameters should be the same for all random variables, in order to get the desired properties.

It is well known that the difference between two independent random variables, that follow a Gumbel distribution, follows a Logistic distribution. It is also well known that the difference between two independent random variables, that follow an Exponential distribution, follow a Laplace distribution. Gumbel is a particular case of ED and Exponential is a particular case of PHRM. Recent research has found similar patterns for the difference between two ED random variables. Srivastava et al. (2006) shows that the difference between two independent ran-

\(^6\)Similar formulations can be used for the other extreme type distributions. Fréchet is a particular case of ED and Weibull of PHRM.
dom variables, that follow an Exponentiated Exponential distribution, follow a Generalized Laplace distribution.

**Lemma 1 (ED Maximum).** In ED, the PDF of the maximum $(n^{th} \text{ OS})$ can be written as

$$f_n(x) = f \left( x, \alpha, \sum_{k=1}^{n} \theta_k \right).$$

**Proof.** The expression from Lemma 1, equation 9, is just a particular case of equation 8 when $r = n$. In that case, the first summation has only one term ($i = 0$). Because $C_n = 1$ and $p_i(n, n) = 1$, the weight of the single PDF is equal to 1. Because $r = n$ and $i = 0$, the second summation also has only one term, and the partial sum inside the density is the total sum that includes all the parameters from $k = 1$ to $r + i = n$. $\square$

We now focus on the first OS (minimum) in Example 1 above. As in the proof of Lemma 1 we just need to follow equation 8. In this case, with $r = 1$, we have three elements in the first summation. The first element is the summation of the partial sum that only includes one random variable. The weights are equal to 1 in this case, and we have three such partial sums ($\theta_1$, $\theta_2$ and $\theta_3$). The second term in the summation is the partial sum that include any two random variables. The weights are equal to -1 in this case and we have three such partial sums ($\theta_1 + \theta_2$, $\theta_1 + \theta_3$ and $\theta_2 + \theta_3$). Because the sign in $p_i(n, r)$ depends on whether $i$ is an even or an odd number, the weights would alternate between a positive and a negative number as we advance in the first summation ($p_0(3, 1) = 1$, $p_1(3, 1) = -1$ and $p_2(3, 1) = 1$). Moreover, for the minimum with $r = 1$ the absolute value of $p_i(n, r)$ will always be 1, therefore the weights alternate between 1 and -1. Finally, as in the case of the maximum, the last term is just one element that includes all random variables ($\theta_1 + \theta_2 + \theta_3$), with weights equal to 1.

Comparing $f_1(x)$ with $f_2(x)$ help us further understand the formulas. First, notice that because $r = 2$ there are no terms which include only one random variable. Second, now the partial sum that includes two random variables ($\theta_1 + \theta_2$, $\theta_1 + \theta_3$ and $\theta_2 + \theta_3$), have positive weights. Finally, notice how the term that includes all random variables ($\theta_1 + \theta_2 + \theta_3$) have a weight that is negative and larger than 1 in absolute value. The result that all weights are equal to 1 in absolute value only applies to the minimum (and trivially to the maximum), but the weights are still all just integers. Finally, notice that in all three order statistics, the weights sum up to 1 ($3 - 3 + 1 = 3 - 2 = 1$). Thus, this example summarizes the nice properties of our results.
Finally, we show the results in the simpler, but much common case in the literature, cases when the distribution are IID or symmetric and when all the random variables come from two groups. In both cases, the formulas are much simpler, because many of the partial sums are now identical, and we do not have to decompose the density over two summations. Lemmas 2 and 3 and below show the results.

**Lemma 2 (ED Symmetric).** If we have \( \theta_1 = \theta_2 = \ldots = \theta_n = \theta \) (the symmetric case), the PDF of OS from ED random variables can be written as

\[
f_r(x) = \sum_{i=0}^{n-r} (-1)^i C_{i-1}^r C_n^t f(x, \alpha, t\theta) .
\]

where \( p_i(n, r) = (-1)^i C_{r+i-1}^i \), \( C_n^t = \frac{n!}{r!(n-r)!} \) and \( t = r + i \).

*Proof.* The expression from Lemma 2, equation 10, is just a particular case of equation 8 when \( \theta_i = \theta \). In that case, each element in the second summation is identical. In other words, each partial summation sums over the same number of \( \theta_i \). In the general case, each \( \theta_i \) is different, so we need to treat them differently in the second summation and in the summation inside the PDF. In the symmetric case, because all \( \theta_i \) are identical, we only need to count. First, we need to count how many elements are in each partial sum, i.e., \( C_n^t \), and we need to count how many \( \theta_i \) are in each partial sum \( t = r + i \). In summary, the results in Lemma 2 are just a particular case of the results in Proposition 1, where the second summation becomes a combinatorial and the summation of parameters is just a multiplication of the symmetric parameter \( \theta \). \( \square \)

Lemma 2 presents results that are much simpler than our general formula in equation 8. The reader can now sum directly over a single summation, rather than having to look over all the subsets in \( S_n \). The reader can check the values in Example 1 using equation 10 for the case with \( \theta_i = \theta \). This would be a particular case of ED with symmetric random variables. In this case, the seven terms in \( f_1(x) \) using equation 8 become only three terms using equation 10. Four of the weights with absolute value of 1 in equation 8 appear in equation 10 with absolute value of 3. In other words, terms that are identical are consolidated with higher weights using equation 10.

**Lemma 3 (ED 2-Groups).** Suppose we have two groups of ED random variables with shape parameters \( \theta_1 \) and \( \theta_2 \). The number of variables in each group is \( n_1 \) and
respectively. Denote \( n = n_1 + n_2 \), we can write the PDF of the OS as

\[
f_r(x) = \sum_{i=0}^{n-r} p_i(n, r) \sum_{k = \max(0, t - n_1)}^{\min(n_2, t)} \binom{n_1}{i-k} \binom{n_2}{k} f(x, \alpha, (t - k)\theta_1 + k\theta_2),
\]

where \( p_i(n, r) = (-1)^i \binom{r+i-1}{r-1}, \binom{n}{r} = \frac{n!}{r!(n-r)!} \) and \( t = r + i \).

Proof. The expression from Lemma 3, equation 11, is just a particular case of equation 8 when \( \theta_i \in \{\theta_1, \theta_2\} \). In that case, several elements in the second summation are identical. In other words, some partial summations sums over the same number subset containing the same numbers of \( \theta_1 \) and \( \theta_2 \). In the general case, each \( \theta_i \) is different, so we need to treat them differently in the second summation and in the summation inside the PDF. In the two-groups case, because there are only two values for \( \theta_i \), we need to count a smaller subset. In summary, the results in Proposition 2 are just a particular case of the results in Proposition 1, where the second summation becomes a combinatorial and the summation of parameters is just a multiplication of the symmetric parameter \( \theta \).

Lemma 3 presents results that are simpler than our general formula in equation 8. The reader can now sum directly over an index \( k \) rather than having to look over all the subsets in \( S_n \). The reader can check the values in Example 1 using equation 11 for the case with \( \theta_3 = \theta_2 \). This would be a particular case of ED with two groups with \( n_1 = 1 \) and \( n_2 = 2 \). In this case, the seven terms in \( f_1(x) \) using equation 8 become only five terms using equation 11. Two of the weights with absolute value of 1 in equation 8 appear in equation 11 with absolute value of 2. In other words, terms that are identical are consolidated with higher weights using equation 11.

### 3.2 Proportional Hazard Rate Models

In this section, we show our results related to Proportional Hazard Rate Models (PHRM) and their properties. Examples of PHRM distributions include Generalized Pareto type II (Lomax), Weibull and Exponential distributions. Proposition 2 shows our main result.

Proposition 2. PHRM Suppose \( X_1, X_2, \ldots, X_n \) are INID PHRM random variables with source CDF \( G(x) \) and shape parameter \( \theta_1, \theta_2, \ldots, \theta_n \). Denote the CDF and PDF of \( X_i \) as \( F(x, \alpha, \theta_i) \) and \( f(x, \alpha, \theta_i) \), which are defined in (3) and (4). Let
$X_{(r)}$ ($r \in \{1, 2, 3, \ldots, n\}$) denote the associated OS in ascending order. The CDF of $X_{(r)}$ can be written as

$$F_r(x) = \sum_{i=0}^{r-1} p_i(n, r) \sum_{S_w} F \left( x, \alpha, \sum_{k=r-i}^{n} \theta_j \right),$$

(12)

The PDF of $X_{(r)}$ can be written as

$$f_r(x) = \sum_{i=0}^{r-1} p_i(n, r) \sum_{S_w} f \left( x, \alpha, \sum_{k=r-i}^{n} \theta_j \right),$$

(13)

where $p_i(n, r) = (-1)^i C_{n+i-r}^i$, $C_n^r = \frac{n!}{r!(n-r)!}$ and $S_w$ extends all combinatorics $(j_{r-i}, j_{r-i+1}, \ldots, j_n)$ of $w = n - r + i + 1$ numbers out of $1, 2, \ldots, n$.

Proof. See Appendix A.3. □

Equations 12 and 13 correspond to equations 5 and 6 in the general case. Notice how the formulas are analogous to equations 7 and 8. The only difference is on the indexes in the summation and the set $S$. In a sense, we are now counting backwards. In the ED case the first summation was from 0 to $n-r$ for partial sums of $r+i$ numbers. In the PHRM, the first summation is from 0 to $r-1$ for partial sums of $n-(r-1)+i$. The results here rely in the similarity of the structure of ED to PHRM. If we use the survival function $H(x)$, i.e., $H(x) = 1 - F(x)$, and apply the definition of PHRM, we get that the survival function of a PHRM is an ED. With this in mind, it is not surprising that the distribution of the maximum of an ED have similar properties than the distribution of the minimum of a PHRM. Moreover, it is easy to extend the results to two additional cases: Inverse Exponentiated Distributions (IED) with $F(x, \theta) = \left[1 - G(x)\right]^\theta$ and Inverse Proportional Hazard Rate Models (IPHRM) with $F(x, \theta) = 1 - G(x)^\theta$. For ease of exposition we will not show the results to these cases. Nonetheless, it is useful to know that the results indeed apply to these four families of distributions: ED, IED, PHRM and IPHRM.

The following corollary shows that the properties of ED, also apply to the case of PHRM.

Corollary 2 (Weights). The weights in the PDF of the OS $X_{(r)}$ from $n$ INID PHRM random variables with shape parameters $\theta_1, \theta_2, \ldots, \theta_n$, do not depend on $\theta_1, \theta_2, \ldots, \theta_n$ and sum up to 1.

Proof. See Appendix A.4. □
We now show an example with $n = 3$ of a general PHRM to reinforce the intuition behind the results, and how they change for PHRM.

**Example 2 (PHRM, $n = 3$).** Suppose we have three PHRM random variables with shape parameters $\theta_1$, $\theta_2$ and $\theta_3$, i.e., $1 - F(x, \theta_i) = [1 - G(x)]^\theta_i$. We can write the PDF of the minimum ($r = 1$), median ($r = 2$) and maximum ($r = 3$) as

$$f_1(x) = f \left( x, \sum_{i=1}^{3} \theta_i \right),$$

$$f_2(x) = f(x, \theta_1 + \theta_2) + f(x, \theta_1 + \theta_3) + f(x, \theta_2 + \theta_3) - 2f \left( x, \sum_{i=1}^{3} \theta_i \right),$$

$$f_3(x) = \sum_{i=1}^{3} f(x, \theta_i) - f(x, \theta_1 + \theta_2) - f(x, \theta_1 + \theta_3) - f(x, \theta_2 + \theta_3) + f \left( x, \sum_{i=1}^{3} \theta_i \right).$$

In the example above, the distribution of the minimum consist of just one distribution. The lemma below shows that this result is general. Moreover, it is the same distribution as that of the maximum in the ED case. Not only that, but the distribution of the median is identical in both cases (when $n$ is odd) and the distribution of the maximum of a PHRM is the same as that of the minimum of an ED. This is a direct result of interpreting the formulas in the PHRM case as counting backwards using the ED formulas. The same intuitions as before apply to this case.

**Lemma 4 (PHRM Minimum).** In PHRM, the PDF of the minimum ($1^{st}$ OS) can be written as

$$f_1(x) = f \left( x, \alpha, \sum_{k=1}^{n} \theta_k \right).$$

**Proof.** The proof is analogous of the proof of Lemma 1. \qed

Finally, as in the previous section, we show the results for the symmetric and the two groups cases.

**Lemma 5 (Symmetric).** In a PHRM with $\theta_1 = \theta_2 = \ldots = \theta_n = \theta$ (symmetry), the OS can be written as

$$f_r(x) = \sum_{i=0}^{r-1} (-1)^i C_{w-1}^i C_{n-w}^{n-w} f(x, \alpha, w\theta).$$
where \( p_i(n, r) = (-1)^i C_{n+i-r}^i \), \( C_n^r = \frac{n!}{r!(n-r)!} \) and \( w = n - r + 1 + i \).

**Proof.** The proof is analogous of the proof of Lemma 2. \( \square \)

**Lemma 6 (PHRM 2-Group).** Suppose we have two groups of random variables with shape parameters \( \theta_1 \) and \( \theta_2 \). The number of variables in each group is \( n_1 \) and \( n_2 \) respectively. Denote \( n = n_1 + n_2 \), we can write the density function of the order statistics as

\[
f_r(x) = \sum_{i=0}^{r-1} p_i(n, r) \sum_{k = \max(0, w-n_1)}^{\min(n_2, w)} C_{n_1}^{w-k} C_{n_2}^k f(x, \alpha, (w - k)\theta_1 + k\theta_2),
\]

where \( p_i(n, r) = (-1)^i C_{n+i-r}^i \), \( C_n^r = \frac{n!}{r!(n-r)!} \) and \( w = n - r + 1 + i \).

**Proof.** The proof is analogous of the proof of Lemma 3. \( \square \)

## 4 Conclusion

We believe our results could be useful in several areas of economics by providing analytical solutions to objects that are commonly used in the literature. The analytical results are both attractive to theorists—who could expand their models and still obtain analytical results—and to econometricians—who could estimate current and future models easily using Maximum Likelihood. An analytical formula for the second order statistic would allow estimation of auctions models to be done using maximum likelihood instead of simulations. Order Statistics from the Gumbel and other Exponentiated Distributions, could be use to estimate discrete-choice models where consumers can buy more than one good. OS from a generalization of the Fréchet and Weibull distributions would allow trade theorists to generalize their current models, while still obtaining an analytical solution to the price distribution. Finally, the results on Proportional Hazard Rates Models could be applied to models with competing risks, allowing for more heterogeneity in the risk types.
A Mathematical Appendix

A.1 Proof of Proposition 1

Before we prove Proposition 1, we now present and prove the following lemma which will be used in the proof of Proposition 1.

**Lemma 7 (Scaled ED).** Suppose \( F_{ji} (i = 1, \ldots, I) \) and \( F_{kl} (l = 1, \ldots, L) \) are ED CDFs with source CDF \( G(x, \alpha) \) and shape parameters \( \theta_{ji} (i = 1, \ldots, I) \) and \( \theta_{kl} (l = 1, \ldots, L) \) respectively. \( f_m \) is the ED PDF with source CDF \( G \) and shape parameter \( \theta_m \). The product \( f_m \prod_{i=1}^{I} F_{ji} \prod_{l=1}^{L} F_{kl} \) is a scaled density of an ED random variable with source CDF \( G \) and shape parameter \( \theta_m + \sum_{i=1}^{I} \theta_{ji} + \sum_{l=1}^{L} \theta_{kl} \).

**Proof.** The proof is straightforward. The product can be expanded to

\[
\begin{align*}
    f_m \prod_{i=1}^{I} F_{ji} \prod_{l=1}^{L} F_{kl} &= \theta_m g(x, \alpha) G(x, \alpha)^{\theta_m - 1} \prod_{i=1}^{I} G(x, \alpha)^{\theta_{ji}} \prod_{l=1}^{L} G(x, \alpha)^{\theta_{kl}} \\
    &= \theta_m g(x, \alpha) G(x, \alpha)^{\theta_m + \sum_{i=1}^{I} \theta_{ji} + \sum_{l=1}^{L} \theta_{kl} - 1} \\
    &= \frac{\theta_m}{\theta^*} f(x, \alpha, \theta^*),
\end{align*}
\]

which is the density of an ED scaled by a constant \( \frac{\theta_m}{\theta^*} \), where \( \theta^* = \theta_m + \sum_{i=1}^{I} \theta_{ji} + \sum_{l=1}^{L} \theta_{kl} \). □

We now present the proof of Proposition 1. **Proof.** The cumulative distribution function (7) can be easily obtained by using the density function, so we will just show how to obtain the density function (8). For simplicity, hereafter we will denote the density function and the cumulative distribution function of a random variable \( X_i \) as \( f_i \) and \( F_i \) respectively. For ease of exposition we would omit the arguments in what follows, e.g., \( f \) and \( F \) (\( f(x) \) and \( F(x) \)) are abbreviations of \( f(x, \alpha, \theta) \) and \( F(x, \alpha, \theta) \).

The probability density function can be derived by first plugging the density function \( f_i \) into the equation 6.

\[
f_r(x) = \frac{1}{(r-1)!(n-r)!} \sum_{S} F_{j_1} \cdots F_{j_{r-1}} f_{j_r} (1 - F_{j_{r+1}}) \cdots (1 - F_{j_n}).
\]

Define \( G_S(x) = (1 - F_{j_{r+1}}) \cdots (1 - F_{j_n}) \) as the probability that the maximum from
the random variables from \( j_{r+1} \) to \( j_n \) would be greater than \( x \). \( S \) extends all the permutations of \((1, 2, \ldots, n)\) and we denote the set all permutations of \((1, 2, \ldots, n)\) by \( P \) i.e. \( P = \{(j_1, \ldots, j_n) : j_i \in \{1, 2, \ldots, n\} \text{ and } j_1 \neq j_2 \neq \ldots \neq j_n\} \). Our proof consists of the following steps

1. Expand \( G_S(x) \) in two cases, \( n - r \) is an odd or an even number.

2. In the summation of each term in the expansion of \( G_S(x) \), we partition \( P \) into subsets. In that way, the summation over \( S \) could be converted to a summation over each subset of \( P \).

3. We then show the summation over each subset of \( P \) produces a weighted density function whose weight does not depend on any parameters.

We now prove the case when \( n - r \) is an odd number. The proof when \( n-r \) is an even number is analogous.

**Case A. \( n - r \) is an odd number.**

In this case, \( G_S(x) \) could be expanded as

\[
G_S(x) = (1 - F_{j_{r+1}}) \ldots (1 - F_{j_n})
\]

\[
= 1 - \sum_{i=r+1}^{n} F_{j_{i}} + \sum_{k=r+1}^{n} \sum_{m=k+1}^{n} F_{j_k} F_{j_m} - \ldots
\]

\[
- F_{j_{r+1}} F_{j_{r+2}} \ldots F_{j_n}.
\]

Define \( L_S(x) = F_{j_{1}} \ldots F_{j_{r-1}} f_{j_r} (1 - F_{j_{r+1}}) \ldots (1 - F_{j_n}) \). We can write

\[
L_S(x) = F_{j_{1}} \ldots F_{j_{r-1}} f_{j_r}
\]

\[
- F_{j_{1}} \ldots F_{j_{r-1}} f_{j_r} \sum_{i=r+1}^{n} F_{j_i} + \ldots
\]

\[
- F_{j_{r+1}} F_{j_{r+2}} \ldots F_{j_n} F_{j_{1}} \ldots F_{j_{r-1}} f_{j_r}.
\]

Applying Lemma 7, the first term is

\[
L_S^1(x) = F_{j_{1}} \ldots F_{j_{r-1}} f_{j_r} = \frac{\theta_{j_r}}{\sum_{i=1}^{r} \theta_{j_i}} f \left( x, \sum_{i=1}^{r} \theta_{j_i} \right).
\]
And the second term is the summation of \( n - r \) similar terms

\[
L_2^n(x) = \sum_{k=r+1}^{n} \frac{\sum_{j=1}^{r} \theta_{j_k}}{\sum_{i=1}^{r} \theta_{j_i}} f \left( x, \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_k} \right).
\]

The third term is a summation over two indexes, and so on,

\[
L_3^n(x) = \sum_{k=r+1}^{n} \sum_{m=k+1}^{n} \frac{\sum_{j=1}^{r} \theta_{j_r} + \theta_{j_k} + \theta_{j_m}}{\sum_{i=1}^{r} \theta_{j_i}} f \left( x, \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_k} + \theta_{j_m} \right).
\]

Finally, the last term could be written as

\[
L_{n-r+1}^n(x) = \frac{\theta_{j_r}}{\sum_{i=1}^{n} \theta_{j_i}} f \left( x, \sum_{i=1}^{n} \theta_{j_i} \right).
\]

Therefore, all terms in \( L_S(x) \) can be expressed as scaled Exponentiated densities. The density function of \( X_{(r)} \) essentially sums over \( S \), i.e., sums over the set \( P \). As a result, \( \sum_S L_S(x) \) is also a summation of scaled Exponentiated densities in which the scale is a function of shape parameters only. We now turn to count the density terms with the same shape parameter in \( \sum_S L_S(x) \) to simplify the density formula.

Recall that \( S \) is all the permutations over \((1, 2, \ldots, n)\) and we have \( n! \) such permutations. Sum over \( S \) for the first term \( L_1^n(x) \) and we have

\[
\sum_S L_1^n(x) = \sum_s \frac{\theta_{j_r}}{\sum_{i=1}^{n} \theta_{j_i}} f \left( x, \sum_{i=1}^{n} \theta_{j_i} \right).
\]

We proceed to calculate the weights of \( f(x, \sum_{i=1}^{r} \theta_{j_i}) \) for each \((j_1, \ldots, j_r)\). We could partition the set \( P \) into \( C_n^r \) mutually exclusive subsets \( P_1, P_2, \ldots, P_{C_n^r} \) sets, i.e., any two of them are exclusive and elements in \( P_k \) have the same first \( r \) values. The total number of such subsets is \( C_n^r \) because we are choosing \( r \) numbers in a given order from \((1, 2, \ldots, n)\). Notice that summing over \( S \) is just summing over all \( P_k \). As each element in \( P_k \) has the same first \( r \) values, their first terms \( \frac{\theta_{j_r}}{\sum_{i=1}^{r} \theta_{j_i}} f(x, \sum_{i=1}^{r} \theta_{j_i}) \) have the same denominators in the weights and same density function \( f \). There are \( r!(n-r)! \) elements in each \( P_k \) and we can partition them again into mutually exclusive sets \( P_k^1, \ldots, P_k^{r} \) such that each element in \( P_k^i \) not only has the same first \( r \) values but also has the exactly same value in the \( r \)th coordinate. Each \( P_k^i \) has cardinality \((r-1)!(n-r)!\). As a result, suppose the first \( r \) values of a permutation
in a particular $P_k$ is $j_1, j_2, ..., j_r$ and the $r$th value of a permutation in a particular $P_k$ is $j_r$, summing over all permutations in $P_k$ would produce

$$\frac{(r-1)!(n-r)!}{\sum_{i=1}^{r} \theta_{j_i}} \sum_{i=1}^{r} \theta_{j_i} f \left(x, \sum_{i=1}^{r} \theta_{j_i}\right)$$

and summing over all $P_k$ of $P_k$ would produce

$$\frac{(r-1)!(n-r)!}{\sum_{i=1}^{r} \theta_{j_i}} \sum_{i=1}^{r} \theta_{j_i} f \left(x, \sum_{i=1}^{r} \theta_{j_i}\right) = (r-1)!(n-r)! \sum_{i=1}^{r} \theta_{j_i} f \left(x, \sum_{i=1}^{r} \theta_{j_i}\right).$$

We then further sum over all $P_k$ which is just a summation of all possible choices of the first $r$ values. Thus

$$\sum_{S} L^1_S(x) = \sum_{P_1, P_2, ..., P_C} L^1_S(x) = (r-1)!(n-r)! \sum_{S_r} f \left(x, \sum_{i=1}^{r} \theta_{j_i}\right),$$

where we denote $S_r$ as extending all combinatorics of $r$ numbers of $1, 2, ..., n$. In the next, we turn the second summation of $L^2_S(x)$ which is expanded as

$$L^2_S(x) = \sum_{k=r+1}^{n} \frac{\theta_{j_r}}{\sum_{i=1}^{r} \theta_{j_i} + \theta_{j_k}} \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_k} f \left(x, \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_k}\right).$$

We will use a similar trick and partition the permutation set $P$ into some subsets in order to easily sum up the weighted densities. We first write $L^2_S(x)$ as

$$L^2_S(x) = \frac{\theta_{j_r}}{\sum_{i=1}^{r} \theta_{j_i} + \theta_{j_r+1}} \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_r+1} f \left(x, \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_r+1}\right) + \frac{\theta_{j_r}}{\sum_{i=1}^{r} \theta_{j_i} + \theta_{j_r+2}} \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_r+2} f \left(x, \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_r+2}\right) + \ldots + \frac{\theta_{j_r}}{\sum_{i=1}^{r} \theta_{j_i} + \theta_{j_n}} \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_n} f \left(x, \sum_{i=1}^{r} \theta_{j_i} + \theta_{j_n}\right),$$

which is a summation of $n-r$ terms. The shape parameter of any density function above depends on $r+1$ parameters. We now first partition the permutation set $P$ into $P_1, P_2, ..., P_C$ and each $P_k$ into $P_1^k, ..., P_r^k$ as defined before. Note that fixing the first $r$ numbers in a permutation is equivalent to fixing the rest $n-r$ numbers. We further partition each $P_k^i$ into $(n-r)!$ subsets $P_k^{i,1}, ..., P_k^{i,(n-r)!}$ such
that each permutation in $P^i_j$ not only has the same $r$ numbers, exactly the same $r$th value, but also has the same permutation of last $n - r$ numbers. Let’s suppose the permutations in a particular $P^i_j$ have the following properties: the first $r$ numbers are $j_1$ to $j_r$, the $r$th value is $j_p$ and the remaining $n - r$ permutations are $(j_{r+1}, \ldots, j_n)$. Thus, $L^2_S(x)$ at each permutation in $P^i_j$ is

$$
\theta_{j_1} \prod_{i=1}^{r-1} \theta_{j_i + \theta_{j_{i+1}}} f \left( x, \sum_{i=1}^{r} \theta_{j_i + \theta_{j_{r+1}}} \right) + \theta_{j_1} \prod_{i=1}^{r-2} \theta_{j_i + \theta_{j_{i+2}}} f \left( x, \sum_{i=1}^{r} \theta_{j_i + \theta_{j_{r+2}}} \right) + \ldots + \\
\theta_{j_1} \prod_{i=1}^{r-1} \theta_{j_i + \theta_{j_n}} f \left( x, \sum_{i=1}^{r} \theta_{j_i + \theta_{j_n}} \right).
$$

Note that the cardinality of $P^i_j$ is $(r - 1)!$. Thus summing over all permutations in $P^i_j$ results in

$$
\frac{(r - 1)! \theta_{j_1}}{\prod_{i=1}^{r-1} \theta_{j_i + \theta_{j_{i+1}}}} f \left( x, \sum_{i=1}^{r} \theta_{j_i + \theta_{j_{r+1}}} \right) + \frac{(r - 1)! \theta_{j_1}}{\prod_{i=1}^{r-2} \theta_{j_i + \theta_{j_{i+2}}}} f \left( x, \sum_{i=1}^{r} \theta_{j_i + \theta_{j_{r+2}}} \right) + \ldots + \\
\frac{(r - 1)! \theta_{j_1}}{\prod_{i=1}^{r-1} \theta_{j_i + \theta_{j_n}}} f \left( x, \sum_{i=1}^{r} \theta_{j_i + \theta_{j_n}} \right).
$$

We denote the above term as $L^2_{P^i_j}(x)$ where $k$ is the subset identity of our first partition of $P$, and $i, j$ are the subset identities of the second and the third partition respectively. Observe that $L^2_{P^i_j}(x)$ is same for all $j$ since the only difference between $L^2_{P^i_j}(x), j = 1, \ldots, (n - r)!$ is just the order of the last $n - r$ terms while
these terms are the same combinatorics. Therefore we have

$$L^2_{P^k}(x) = \sum_{j=1}^{(n-r)!} \sum_{i=1}^{r} \frac{(n-r)!}{\hat{\theta}_{j_i} + \hat{\theta}_{j_{r+1}}} f \left( x, \sum_{i=1}^{r} \hat{\theta}_{j_i} + \hat{\theta}_{j_{r+1}} \right) + \frac{(n-r)!}{\hat{\theta}_{j_i} + \hat{\theta}_{j_{r+2}}} f \left( x, \sum_{i=1}^{r} \hat{\theta}_{j_i} + \hat{\theta}_{j_{r+2}} \right) + \ldots + \frac{(n-r)!}{\hat{\theta}_{j_i} + \hat{\theta}_{j_{n}}} f \left( x, \sum_{i=1}^{r} \hat{\theta}_{j_i} + \hat{\theta}_{j_{n}} \right)$$

= \sum_{m=r+1}^{n} \frac{(n-r)!}{\hat{\theta}_{j_i} + \hat{\theta}_{j_m}} f \left( x, \sum_{i=1}^{r} \hat{\theta}_{j_i} + \hat{\theta}_{j_m} \right) \cdot \sum_{i=1}^{r} \hat{\theta}_{j_i} + \hat{\theta}_{j_m}.$$

We then sum all subsets of $P_k$, in which the first $r$ values are same (thus are the last $n-r$ values).

$$L^2_{P^k}(x) = \sum_{p=1}^{r} \sum_{m=r+1}^{n} \frac{(n-r)!}{\hat{\theta}_{j_i} + \hat{\theta}_{j_m}} f \left( x, \sum_{i=1}^{r} \hat{\theta}_{j_i} + \hat{\theta}_{j_m} \right)$$

= \frac{(n-r)!}{\hat{\theta}_{j_i} + \hat{\theta}_{j_{r+1}}} f \left( x, \sum_{i=1}^{r} \hat{\theta}_{j_i} + \hat{\theta}_{j_{r+1}} \right) + \frac{(n-r)!}{\hat{\theta}_{j_i} + \hat{\theta}_{j_{r+2}}} f \left( x, \sum_{i=1}^{r} \hat{\theta}_{j_i} + \hat{\theta}_{j_{r+2}} \right) + \ldots + \frac{(n-r)!}{\hat{\theta}_{j_i} + \hat{\theta}_{j_{n}}} f \left( x, \sum_{i=1}^{r} \hat{\theta}_{j_i} + \hat{\theta}_{j_{n}} \right)$$

Notice that we can interpret the above summation as fixing $(j_{r+1}, \ldots, j_n)$. Due to the way we construct it, the exact permutation $(j_{r+1}, \ldots, j_n)$ we use is irrelevant. We begin with a particular $P^i_k$ where the last $n-r$ numbers have a particular permutation. All possible permutations of the last $n-r$ values will result in the same $L^2_{P^k}(x)$, as long as the first $r$ numbers are fixed in $P_k$. Consequently, we read $L^2_{P^k}(x)$ for each $k$ as fixing the first $r$ numbers and fixing one particular permutation of the last $n-r$ numbers. Additionally, the weights of the density functions in $L^2_{P^k}(x)$ differ only in one shape parameter in the denominator.

At this step, we have expressed $L^2_{P^k}(x)$ as a summation of $n-r$ weighted density functions. Therefore $\sum_S L^2_S(x)$ is reduced to a summation of $(n-r)C^r_n$ weighted density functions. Observe that the denominator in each weight is a summation of $r+1$ terms, while we only partition $P$ by the first $r$ values in the permutation.
Therefore, two different weights of the density function in two different subsets, \(P_i\) and \(P_j\) where \(i \neq j\), could have the same denominators which encourage us to further simplify the second term in \(\sum_S L_S(x)\). The next step is to take out the weights and sum them up given the same parameters in a weighted density function.

Given \(r + 1\) parameters from \(\theta_1, \theta_2, ..., \theta_n\), we consider the weighted density functions which have the shape parameter as their summation in \(\sum_S L_S^2(x)\). Let’s consider a given set \(E = \{\theta_{m_1}, ..., \theta_{m_{r+1}}\}\) and think which density functions in our summation have the shape parameters as the sum of all numbers in \(E\). It is easy to see that, as long as the first \(r\) numbers in a permutation for a set \(P_k\) fall in the set \(E\), there must be such a weighted density function. In each summation over \(P_k\), there is only one weighted density function, the denominator of whose weight is \(\sum_{i=1}^{r+1} \theta_{m_i}\). Collect all the weights and we could get

\[
(n - r)!(r - 1)! \frac{\sum_{i=1,i\neq 1}^{r+1} \theta_{m_i} f \left( x, \sum_{i=1}^{r+1} \theta_{m_i} \right)}{\sum_{i=1}^{r+1} \theta_{m_i}} \\
(n - r)!(r - 1)! \frac{\sum_{i=1,i\neq 2}^{r+1} \theta_{m_i} f \left( x, \sum_{i=1}^{r+1} \theta_{m_i} \right)}{\sum_{i=1}^{r+1} \theta_{m_i}} \\
\vdots \\
(n - r)!(r - 1)! \frac{\sum_{i=1,i\neq r+1}^{r+1} \theta_{m_i} f \left( x, \sum_{i=1}^{r+1} \theta_{m_i} \right)}{\sum_{i=1}^{r+1} \theta_{m_i}} \]

which are \(r + 1\) terms in total and each of them comes from the summation over one \(P_k\) subset. Summing all terms above, we have

\[
(n - r)!(r - 1)! \frac{\sum_{i=1}^{r+1} \theta_{m_i} f \left( x, \sum_{i=1}^{r+1} \theta_{m_i} \right)}{\sum_{i=1}^{r+1} \theta_{m_i}} = r(n - r)!(r - 1)! f \left( x, \sum_{i=1}^{r+1} \theta_{m_i} \right). 
\]

Thus \(r + 1\) terms are reduced to one single term. We repeat this procedure for each possible set \(E\) i.e. for each \(r + 1\) combination from 1 to \(n\). We then reduce \(\sum_S L_S^2(x)\) to a summation of terms as above over different \(\theta_{m_1}, ..., \theta_{m_{r+1}}\) which is

\[
\sum_{S_{r+1}} r(n - r)!(r - 1)! f \left( x, \sum_{i=1}^{r+1} \theta_{m_i} \right),
\]

where \(S_{r+1}\) is represented as extending all combinatorics of \(r + 1\) numbers of 1, 2, ..., \(n\). The above formula has \(\binom{n}{r+1}\) terms, which is consistent with our re-
duction of \((n - r)C^n_r\) terms by summing a group of \((r + 1)\) of them together, i.e. \\
\[\frac{(n-r)C^n_r}{r+1} = C_{r+1}^n.\]

Consequently, when \(n-r\) is an odd number, we could write \(f_r(x)\) in the following way
\[
f_r(x) = \frac{1}{(r-1)!(n-r)!} \sum_{S} L_S(x) = \frac{1}{(r-1)!(n-r)!}[L^1_S(x) - L^2_S(x) + \ldots - L^{n-r+1}_S(x)]
\]
\[
= \sum_{S} \sum_{i=1}^{r} f \left( x, \sum_{j=1}^{r} \theta_{j,i} \right) - \sum_{S_{r+1}} r f \left( x, \sum_{i=1}^{r+1} \theta_{j,i} \right) + \ldots - \sum_{S_n} \frac{(n-1)!}{(n-r)!(r-1)!} f \left( x, \sum_{i=1}^{n} \theta_i \right)
\]
\[
= \sum_{i=0}^{n-r} p_i(n, r) \sum_{S_{r+i}} f \left( x, \sum_{j=1}^{r+i} \theta_{j,k} \right)
\]
\[
= \sum_{i=0}^{n-r} p_i(n, r) \sum_{S_{r+i}} f \left( x, \sum_{j=1}^{r+i} \theta_{j,k} \right)
\]

**Case B. \(n - r\) is an even number.**

When \(n - r\) is an even number, we extend \(G_S(x)\) in the following way
\[
G_S(x) = (1 - F_{j_{r+1}}) \ldots (1 - F_{j_{n}})
\]
\[
= 1 - \sum_{i=r+1}^{n} F_i + \sum_{k=r+1}^{n} \sum_{m=k+1}^{n} F_{j_k} F_{j_m} - \ldots
\]
\[
\ldots + F_{j_{r+1}} F_{j_{r+2}} \ldots F_{j_n},
\]

The rest part to prove the result is analogous as that when \(n - r\) is an odd number.

Thus we have proved that the density function of \(X_{(r)}\) in Proposition (1) takes the form (8). □
A.2 Proof of Corollary 1

Proof. For each $i$ from 0 to $n - r$, there are in total $C_{n}^{r+i}$ combinatorics for a given order $r$. Summing up all weights of the density functions $f(x, \alpha, \cdot)$ results in

$$W = \sum_{i=0}^{n-r} p_i(n, r) C_{n}^{r+i} = \sum_{i=0}^{n-r} (-1)^{i} C_{r+i-1}^{r+i} C_{r+i}^{n}$$

$$= \sum_{i=0}^{n-r} (-1)^{i} \frac{n!}{(r+i)! (n-r-i)!}$$

$$= \sum_{i=0}^{n-r} \frac{(n-r)!}{i! (n-r-i)!} \frac{n!}{(r+i)! (n-r-i)!}$$

$$= \sum_{i=0}^{n-r} \frac{(-1)^{i} C_{n-r}^{r} C_{n}^{r+i}}{r+i} = 1,$$

where the last identity follows formula (1.41) in Gould (1972) when we use $n^* = n - r$, $k = i$, $x = r$ and notice that $C_{n}^{r} = C_{n}^{n-r}$. □

A.3 Proof of Proposition 2

Before we prove Proposition 2, we now present and prove the following lemma which will be used in the proof of Proposition 2.

Lemma 8 (Scaled PHRM). Suppose $F_{j_i}$ ($i = 1, .., I$) and $F_{k_l}$ ($l = 1, .., L$) are reversed Exponentiated CDFs with source CDF $G(x, \alpha)$ and shape parameters $\theta_{j_i}$ ($i = 1, .., I$) and $\theta_{k_l}$ ($l = 1, .., L$) respectively. $f_m$ is the reversed Exponentiated PDF with source CDF $G$ and shape parameter $\theta_m$. The product $f_m \prod_{i=1}^{I} (1-F_{j_i}) \prod_{l=1}^{L} (1-F_{k_l})$ is a scaled density of a reversed Exponentiated random variable with source CDF $G$ and shape parameter $\sum_{i=1}^{I} \theta_{j_i} + \sum_{l=1}^{L} \theta_{k_l} + \theta_m$.

Proof. The product can be expanded to

$$f_m \prod_{i=1}^{I} (1-F_{j_i}) \prod_{l=1}^{L} (1-F_{k_l})$$

$$= \theta_m g(x, \alpha)[1 - G(x, \alpha)]^{\theta_m - 1} \prod_{i=1}^{I} [1 - G(x, \alpha)]^{\theta_{j_i}} \prod_{l=1}^{L} [1 - G(x, \alpha)]^{\theta_{k_l}}$$

$$= \theta_m g(x, \alpha)[1 - G(x, \alpha)]^{\theta_m + \sum_{i=1}^{I} \theta_{j_i} + \sum_{l=1}^{L} \theta_{k_l} - 1}$$

$$= \frac{\theta_m}{\theta^*} f(x, \alpha, \theta^*)$$

22
which is the density of a reversed Exponentiated distribution scaled by a constant \( \frac{\theta_m}{\theta^*} \), where \( \theta^* = \theta_m + \sum_{i=1}^{I} \theta_{j_i} + \sum_{i=1}^{L} \theta_{k_i}. \)

We now present the proof for Proposition 2. Notice that the formulation for PHRM is very similar to that of ED. In particular, it correspond to the same formulation if we use the survival function, instead of the PDF. The proof for PHRM is analogous to the proof for ED when we work with survival functions rather than PDF, and we relabel the indexes. Therefore, we will only provide a concise proof here since the strategy we use to prove Proposition 2 is essentially the counting method in proving Proposition 1. Proof. Denote the survival function of \( F(x, \alpha, \theta_i) \) by \( H_i = 1 - F(x, \alpha, \theta_i) \).

Recall that equation 6 states that

\[
 f_r(x) = \frac{1}{(r-1)!(n-r)!} \sum_{S} F_{j_1} \cdots F_{j_{r-1}} f_{j_r}(1 - F_{j_{r+1}}) \cdots (1 - F_{j_n})
\]

Equivalently, using our new notations we could write

\[
 f_r(x) = \frac{1}{(r-1)!(n-r)!} \sum_{S} (1 - H_{j_1}) \cdots (1 - H_{j_{r-1}}) f_{j_r} H_{j_{r+1}} \cdots H_{j_n}
\]

Alternatively we define the product of the first \( r - 1 \) terms as \( G_S(x) = (1 - H_{j_1}) \cdots (1 - H_{j_{r-1}}) \) and distinguish two cases, \( r - 1 \) is an odd number and \( r - 1 \) is an even number, to find the density function. Let’s consider the case when \( r - 1 \) is an even number. In this case, we could write

\[
 G_S(x) = 1 - \sum_{i=1}^{r-1} H_{j_i} + \sum_{k=1}^{r-1} \sum_{m=k+1}^{r-1} H_{j_k} H_{j_m} - \cdots - H_{j_1} H_{j_2} \cdots H_{j_{r-1}}.
\]

Similarly we write

\[
 L_S(x) = G_S(x) f_{j_r} H_{j_{r+1}} \cdots H_{j_n}
\]

\[
 = f_{j_r} H_{j_{r+1}} \cdots H_{j_n} - f_{j_r} H_{j_{r+1}} \cdots H_{j_n} \sum_{i=1}^{r-1} H_{j_i} + \cdots - H_{j_1} H_{j_2} \cdots H_{j_{r-1}} f_{j_r} H_{j_{r+1}} \cdots H_{j_n}.
\]

We denote the i-th term in \( L_S(s) \) by \( L_S^i(s) \), then the density function could be
written as
\[
f_r(x) = \frac{1}{(r-1)!(n-r)!} \sum_S [L_1^S(x) - L_2^S(x) + \ldots - L_r^S(x)].
\]

Using the same counting method we proposed in the proof of Proposition 1 and Lemma 8, we know that

\[
\sum_S L_1^S(x) = \sum_S \frac{\theta_{j_r}}{\sum_{i=r}^n \theta_{j_i}} f \left( x, \sum_{i=r}^n \theta_{j_i} \right)
= \sum_{S_{n-r+1}} \frac{n!}{(n-r+1)C_{n-1}^r} f \left( x, \sum_{i=r}^n \theta_{j_i} \right).
\]

And

\[
L_2^S(x) = \frac{\theta_{j_r}}{\sum_{i=r}^n \theta_{j_i} + \theta_{j_{r-1}}} f \left( x, \sum_{i=r}^n \theta_{j_i} + \theta_{j_{r-1}} \right) + \ldots
= \frac{\theta_{j_r}}{\sum_{i=r}^n \theta_{j_i} + \theta_{j_{r-2}}} f \left( x, \sum_{i=r}^n \theta_{j_i} + \theta_{j_{r-2}} \right) + \ldots
= \frac{\theta_{j_r}}{\sum_{i=r}^n \theta_{j_i} + \theta_{j_1}} f \left( x, \sum_{i=r}^n \theta_{j_i} + \theta_{j_1} \right).
\]

Similarly, we have

\[
\sum_S L_3^S(x) = \sum_{S_{n-r+2}} \frac{n!(r-1)}{(n-r+2)C_n^{r-2}} f \left( x, \sum_{i=r-1}^n \theta_{j_i} \right).
\]
Finally, we get
\[
f_r(x) = \frac{1}{(r-1)!(n-r)!} \sum_s [L_s^1(x) - L_s^2(x) + \ldots - L_s^n(x)]
\]
\[
= \sum_{s_{n-r+1}} f\left(x, \sum_{i=r}^n \theta_j\right) - \sum_{s_{n-r+2}} (n-r+1) f\left(x, \sum_{i=r-1}^n \theta_j\right) + \ldots - \frac{(n-1)!}{(r-1)!(n-r)!} f\left(x, \sum_{i=1}^n \theta_j\right)
\]
\[
= \sum_{i=0}^{r-1} (-1)^i C_{n+i-r}^i \sum_{s_{n-r+1+i}} f\left(x, \sum_{k=r-i}^n \theta_j\right)
\]
\[
= \sum_{i=0}^{r-1} (-1)^i C_{n+i-r}^i \sum_{s_{n-r+1+i}} f\left(x, \sum_{k=r-i}^n \theta_j\right)
\]

When \( r-1 \) is an odd number, we can extend \( G_S(x) \) as follows
\[
G_S(x) = 1 - \sum_{i=1}^{r-1} H_j + \sum_{k=1}^{r-1} \sum_{m=k+1}^{r-1} H_j H_{jm} - \ldots + H_{j_1} H_{j_2} \ldots H_{j_{r-1}}
\]
and the rest of the proof is similar as when \( r-1 \) is an even number. \( \square \)

A.4 Proof of Corollary 2

Proof. For each \( i \) from 0 to \( r-1 \), there are in total \( C_{n-r+i+1}^{n-r+i-1} = C_{n-r+i-1}^{r-i-1} \) combinatorics for a given order \( r-i-1 \). Summing up all weights of the density functions \( f(x, \alpha, \cdot) \) results in
\[
W = \sum_{i=0}^{r-1} p_i(n, r) C_{n-r+i-1}^{r-i-1} = \sum_{i=0}^{r-1} (-1)^i C_{n+i-r}^i C_{n-r}^{r-i-1}
\]
\[
= \sum_{i=0}^{r-1} (-1)^i \frac{n!}{(n-r+i+1)! (n-r)!(r-1-i)!}
\]
\[
= \sum_{i=0}^{r-1} (-1)^i \frac{(r-1)!}{i!(r-1-i)!} \frac{n!}{(n-r+1)! (n-r+i+1)!}
\]
\[
= \sum_{i=1}^{r-1} (-1)^i C_{r-1}^i C_{n-r+i}^{n-r+1} = 1,
\]
where the last identity follows formula (1.41) in Gould (1972), when we use \( n^* = \ldots \)
$r - 1, \ k = i, \ x = n - r + 1$ and notice that $C^r_n = C^{n-r}_n$.
References


