A general noncentral hypergeometric distribution

Simon Loertscher[1], Ellen V. Muir[2] and Peter G. Taylor[2]

[1]Department of Economics, The University of Melbourne

[2]School of Mathematics and Statistics, The University of Melbourne

May 25, 2016

Abstract

We construct a general noncentral hypergeometric distribution, which models biased sampling without replacement. Our distribution is constructed from the combined order statistics of two samples; one of independent and identically distributed random variables with absolutely continuous distribution F and the other of independent and identically distributed random variables with absolutely continuous distribution G. The distribution depends on F and G only through $F \circ G^{(-1)}$ (F composed with the quantile function of G) and the standard hypergeometric distribution and Wallenius' noncentral hypergeometric distribution arise as special cases. We show in efficient economic markets the quantity traded has a general noncentral hypergeometric distribution.

Keywords Noncentral hypergeometric distributions, Wallenius, sampling without replacement, two-sample order statistics, mechanism design, efficient market quantity.

1 Introduction

The standard hypergeometric distribution models the number of marked objects obtained when an unbiased sample is collected from a finite population without replacement. Fisher's and Wallenius' noncentral hypergeometric distributions are widely used generalisations of this standard distribution. Fisher's noncentral hypergeometric distribution was first described by Fisher (1935) in the context of contingency tables. Wallenius (1963) conceived a noncentral hypergeometric distribution in his PhD thesis (later published as a technical report by Stanford University) for use in competitive models of Darwinian evolution. These noncentral hypergeometric distributions may be considered in the context of an urn problem with simple biased sampling. Specifically, sampling is undertaken such that if an urn contains balls of two different weights, the probability of drawing a given ball is proportional to its weight.

The difference between the Fisher and Wallenius distributions is subtle and has led to nomenclature confusion, which was resolved by Fog (2008). Fisher's distribution models a situation in which balls are sampled independently and Wallenius' distribution models

Address correspondence to Ellen V. Muir, School of Mathematics and Statistics, The University of Melbourne, Parkville VIC 3010, Australia; Email: evmuir@unimelb.edu.au

sampling without replacement. However, when sampling is unbiased — that is, all balls are of equal weight — both distributions give rise to the standard hypergeometric distribution. This is discussed in Section 2 in more detail.

In this paper we show it is possible to consider the standard hypergeometric distribution in terms of samples taken from two independent sets of order statistics drawn from the same distribution. It is natural to extend this to samples of sets of order statistics drawn from different distributions, F and G. This allows us to construct a general noncentral hypergeometric distribution which models cases in which marked objects are more or less likely to be sampled. Let $G^{(-1)}$ denote the quantile function of G (which in this case is simply the inverse of G) and denote by G^ω the function $G^\omega(x) = (G(x))^\omega$. We will show our noncentral hypergeometric distribution depends on the functions F and G only through $F \circ G^{(-1)}$. In some robust special cases, we obtain a distribution which depends on F and G through a small number of parameters. In particular, when $G = F^\omega$, it depends only on ω.

The remainder of the paper is organised as follows. In Section 2 we discuss the three well-known existing hypergeometric distributions. Section 3 provides a construction of our general noncentral hypergeometric distribution using the combined order statistics of two samples, one comprised of independent and identically distributed random variables with an absolutely continuous distribution F. For the other, the absolutely continuous distribution function is G. In Section 4, our noncentral hypergeometric distribution is explicitly computed in terms of the functions F and G.

Section 5 shows the standard hypergeometric distribution and Wallenius' noncentral hypergeometric distribution arise as special cases. In Section 6 we show the general noncentral hypergeometric distribution depends on F and G through $F \circ G^{(-1)}$. Finally, in Section 7 we present an example which shows the general noncentral hypergeometric distribution arises naturally in the study of economic markets. In particular, the efficient quantity traded in the canonical mechanism design setup has our general noncentral hypergeometric distribution. We also present an application for our result regarding Wallenius' noncentral hypergeometric distribution. Some concluding remarks are contained in Section 8, with lengthy proofs included in subsequent appendices.

2 Hypergeometric Distributions

2.1 The Standard Hypergeometric Distribution

Definition 2.1. *Let $N \in \mathbb{N}$, $n \in \{0, 1, \ldots, N\}$ and $D \in \{0, 1, \ldots, N\}$ be given. If, for $x = \max\{0, n + D - N\}, \ldots, \min\{n, D\}$, X has a probability mass function given by*

$$p_X(x; n, N, D) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}}, \tag{1}$$

then we say X has a standard hypergeometric distribution *and write $X \stackrel{d}{=} Hg(n, D, N)$.*

Suppose an urn contains N balls, D of which are marked. If n successive draws are performed without replacement, the number of marked balls contained within the sample has a standard hypergeometric distribution.

2.2 Noncentral Hypergeometric Distributions

Definition 2.2. *Let $N \in \mathbb{N}$, $n \in \{0, 1, \ldots, N\}$, $D \in \{0, 1, \ldots, N\}$ and $\omega_1, \omega_2 \in (0, \infty)$ be given, with $\omega = \omega_1/\omega_2$. Suppose for $x = \max\{0, n + D - N\}, \ldots, \min\{n, D\}$, X has a probability mass function given by*

$$p_X(x; n, N, D, \omega) = \frac{\binom{D}{x}\binom{N-D}{n-x}\omega^x}{P_0}, \quad \textit{where} \quad P_0 = \sum_{y=\max\{0,n+D-N\}}^{\min\{n,D\}} \binom{D}{y}\binom{N-D}{n-y}\omega^y. \tag{2}$$

Then we write $X \stackrel{d}{=} HgF(n, D, N, \omega)$ and say X has Fisher's noncentral hypergeometric distribution.

Fisher's distribution may be described in the context of an urn problem in the following manner. Suppose an urn contains N balls and D of these balls are marked. A sample of Y_1 marked balls is collected by including each ball with probability $\omega_1/(\omega_1+\omega_2)$. A sample of Y_2 unmarked balls is then collected by including each ball with probability $\omega_2/(\omega_1+\omega_2)$. With this method of sampling, the event that one ball is included in the sample is independent of the inclusion of other balls. Fisher's noncentral hypergeometric distribution is given by the number of marked balls Y_1, conditional on $Y_1 + Y_2 = n$. Thus, Fisher's distribution is the conditional distribution of two independent binomial random variables, given their sum.

If $\omega_1 = \omega_2$, marked and unmarked balls are equally likely to be included in the sample. In this case, Fisher's noncentral hypergeometric distribution is equivalent to the standard hypergeometric distribution.

Definition 2.3. *Let* $N \in \mathbb{N}$, $n \in \{0, 1, \ldots, N\}$, $D \in \{0, 1, \ldots, N\}$ *and* $\omega_1, \omega_2 \in (0, \infty)$ *be given, with* $\omega = \omega_1/\omega_2$. *Suppose for* $x = \max\{0, n + D - N\}, \ldots, \min\{n, D\}$, X *has a probability mass function given by*

$$p_X(x; n, N, D, \omega) = \binom{D}{x}\binom{N-D}{n-x}\int_0^1 (1 - t^{1/d})^{n-x}(1 - t^{\omega/d})^x \, dt, \tag{3}$$

where $d = x - D + N - n + \omega(D - x)$. *Then* X *has* Wallenius' noncentral hypergeometric distribution *and we write* $X \stackrel{d}{=} HgW(n, D, N, \omega)$.

Wallenius' noncentral hypergeometric distribution may also be described in terms of sampling without replacement. Consider an urn which contains D marked balls of weight ω_1 and $N - D$ unmarked balls of weight ω_2. Assume n successive draws are performed such that the probability of selecting a given ball is its proportion of the total weight of all remaining balls. Then the number of marked balls within the sample of size n has Wallenius' noncentral hypergeometric distribution.

Wallenius (1963) showed that the probability mass function of the number of marked balls selected satisfies the combinatorial recursion

$$\begin{aligned} p_X(x; n, N, D, \omega) = {} & \frac{\omega(D - x + 1)}{\omega(D - x + 1) + N - D - n + x} \, p_X(x - 1; n - 1, N, D, \omega) \\ & + \frac{N - D - n + x + 1}{\omega(D - x) + N - D - n + x + 1} \, p_X(x, n - 1, N, D, \omega) \end{aligned} \tag{4}$$

and the solution to this recursion is given by (3).

If the sampling is performed without bias, then marked and unmarked balls have an equal weight. Thus, the standard hypergeometric distribution is obtained as a special case of Wallenius' noncentral hypergeometric distribution when $\omega_1 = \omega_2$.

We have obtained the standard noncentral hypergeometric distribution as a special case of both Fisher's and Wallenius' distributions. However, Wallenius' noncentral hypergeometric distribution is the natural generalisation of the standard hypergeometric distribution, as it models sampling without replacement. When balls are sampled without replacement from an urn, it is necessarily the case that successive draws are dependent. Since Fisher's noncentral hypergeometric distribution is constructed such that individual draws are independent, no method of sampling without replacement will give rise to this distribution in general, with the notable exception being the case in which $\omega_1 = \omega_2$.

The remainder of this paper will focus on the construction of a general noncentral hypergeometric distribution, which models biased sampling without replacement. We do not expect Fisher's noncentral hypergeometric distribution to arise naturally as a special case, as this distribution does not model sampling without replacement, in general.

3 Order Statistics and Sampling Without Replacement

In this section, we show it is possible to consider sampling without replacement in terms of the order statistics of two samples of random variables. For ease of exposition, we consider absolutely continuous probability distribution functions only. However, the results presented in this paper can be generalised to Riemann-Stieltjes integrable functions.

Take $N \in \mathbb{N}$, $n \in \{0, 1, \ldots, N\}$ and $D \in \{0, 1, \ldots, N\}$ and let F and G be absolutely continuous probability distribution functions on $\mathbb{R}$. Let $\mathcal{X} = \{X_1, \ldots, X_D\}$ be a sample of independent and identically distributed random variables with distribution F. Similarly, let $\mathcal{Y} = \{Y_1, \ldots, Y_{N-D}\}$ be a sample of independent and identically distributed random variables with distribution G. For ease of exposition, we assume the random variables in the set $\mathcal{Z} = \mathcal{X} \cup \mathcal{Y}$ are distinct (an event which occurs almost surely since F and G are absolutely continuous).

Order the random variables in the set $\mathcal{Z}$ from smallest to largest.[1] We consider the first n random variables from the ordered list (the n smallest random variables from the set $\mathcal{Z}$) and count the number of random variables from the set $\mathcal{X}$ within this sample. Denote this random variable by Q. Here, Q lies within the set $\{\max\{0, n+D-N\}, \ldots, \min\{n, D\}\}$.

Formally, let $Z_1 = X_1, \ldots, Z_D = X_D, Z_{D+1} = Y_1, \ldots, Z_N = Y_{N-D}$. Using indicator functions, Q may be written

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{D} \mathbb{I}(Z_{(i)} = X_j), \tag{5}$$

where $Z_{(i)}$ is the ith order statistic, that is, the ith lowest random variable in the set $\mathcal{Z}$. Refer to David (1970) for an introduction to order statistics.

By construction, Q models biased sampling without replacement. Thus, Q has a noncentral hypergeometric distribution. However, the nature of the biased sampling depends non-trivially on the distribution functions F and G. We thus have defined a general noncentral hypergeometric distribution, dependent on n, D, N, F and G.

[1] If F and G are Riemann-Stieltjes integrable the random variables in the set $\mathcal{Z}$ are not necessarily distinct and a strict ordering of the set is required to proceed. One possibility is to resolve ties randomly and uniformly. Once the set $\mathcal{Z}$ has been strictly ordered, all results presented in the paper generalise.

We next show it is possible to define Q in terms of the order statistics of the sets $\mathcal{X}$ and $\mathcal{Y}$. Let $Y_{[j]}$ denote the jth largest random variable from the set $\mathcal{Y}$.

We start by proving the following lemma. Let $\underline{k} = \max\{0, n+D-N\}$ and $\overline{k} = \min\{n, D\}$ so the support of Q becomes $\{\underline{k}, \ldots, \overline{k}\}$ for the remainder of the paper.

Lemma 3.1. *Take $k \in \{\underline{k}+1, \ldots, \overline{k}\}$. Then $X_{(k)} < Y_{[N-D-n+k]}$ if and only if there are at least k random variables from $\mathcal{X}$ among the n smallest random variables from $\mathcal{Z}$. That is, $X_{(k)} < Y_{[N-D-n+k]}$ if and only if*

$$\sum_{i=1}^{n}\sum_{j=1}^{D} \mathbb{I}(Z_{(i)} = X_j) \geq k. \tag{6}$$

Proof. Suppose (6) holds and there are at least k random variables from $\mathcal{X}$ among the n smallest random variables from $\mathcal{Z}$. Then there are at least $N-D-n+k$ random variables from $\mathcal{Y}$ among the $N-n$ largest random variables from $\mathcal{Z}$. Thus, $X_{(k)} < Y_{[N-D-n+k]}$. Conversely, suppose $X_{(k)} < Y_{[N-D-n+k]}$. Then exactly $D-k$ random variables from $\mathcal{X}$ are greater than $X_{(k)}$ and at least $N-D-n+k$ random variables from $\mathcal{Y}$ are greater than $X_{(k)}$. Therefore, at least $N-n$ random variables from $\mathcal{Z}$ are greater than $X_{(k)}$ and $X_{(k)}$ is contained in the lowest n random variables from $\mathcal{Z}$. Thus, the lowest n random variables from $\mathcal{Z}$ contains at least k random variables from $\mathcal{X}$ and (6) holds. □

Using Lemma 3.1 we prove a theorem expressing Q in terms of order statistics.

Theorem 3.1. *1. $Q = \underline{k}$ if and only if $Y_{[N-D-n+\underline{k}+1]} < X_{(\underline{k}+1)}$*

2. For $\underline{k} < k < \overline{k}$, $Q = k$ if and only if $X_{(k)} < Y_{[N-D-n+k]}$ and $Y_{[N-D-n+k+1]} < X_{(k+1)}$

3. $Q = \overline{k}$ if and only if $X_{(\overline{k})} < Y_{[N-D-n+\overline{k}]}$

Proof. Suppose $\underline{k} < k < \overline{k}$. By Lemma 3.1

$$\sum_{i=1}^{n}\sum_{j=1}^{D} \mathbb{I}(Z_{(i)} = X_j) \geq k \tag{7}$$

if and only if $X_{(k)} < Y_{[N-D-n+k]}$ and

$$\sum_{i=1}^{n}\sum_{j=1}^{D} \mathbb{I}(Z_{(i)} = X_j) < k+1 \tag{8}$$

if and only if $Y_{[N-D-n+k+1]} < X_{(k+1)}$ (since $Y_{[N-D-n+k+1]} \neq X_{(k+1)}$ by assumption). Combining (7) and (8) shows

$$\sum_{i=1}^{n}\sum_{j=1}^{D} \mathbb{I}(Z_{(i)} = X_j) = k \tag{9}$$

if and only if $X_{(k)} < Y_{[N-D-n+k]}$ and $Y_{[N-D-n+k+1]} < X_{(k+1)}$. So by (9) and the definition of Q, $Q = k$ if and only if $X_{(k)} < Y_{[N-D-n+k]}$ and $Y_{[N-D-n+k+1]} < X_{(k+1)}$.

The $Q = \underline{k}$ and $Q = \overline{k}$ cases are similar. □

4 The Probability Mass Function of Q

As seen in Section 3, Q has a noncentral hypergeometric distribution which depends non-trivially on the functions F and G. This gives rise to a general noncentral hypergeometric distribution. Using Theorem 3.1, the probability mass function of Q can be computed.

When $\underline{k} = 0$ we have

$$\mathbf{P}(Q = \underline{k}) = \frac{(N-D)!}{(N-D-n)!(n-1)!} \int_{-\infty}^{\infty} [1 - F(s)]^D G^{n-1}(s)[1 - G(s)]^{N-D-n} g(s)\, ds, \quad (10)$$

and when $\underline{k} = n + D - N$,

$$\mathbf{P}(Q = \underline{k}) = \tfrac{D!}{(n+D-N)!(N-n-1)!} \int_{-\infty}^{\infty} F^{n+D-N}(s)[1 - F(s)]^{N-n-1} f(s) G^{N-D}(s)\, ds. \quad (11)$$

For the case in which $\underline{k} < k < \overline{k}$,

$$\begin{aligned}
&\mathbf{P}(Q = k) \\
&= \tfrac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!} \Bigg(\int_{-\infty}^{\infty} F^k(s)[1 - F(s)]^{D-k-1} f(s) G^{n-k}(s)[1 - G(s)]^{N-D-n+k}\, ds \\
&\quad + \tfrac{N-D-n+k}{D-k} \int_{-\infty}^{\infty} F^k(s)[1 - F(s)]^{D-k} G^{n-k}(s)[1 - G(s)]^{N-D-n+k-1} g(s)\, ds \Bigg). \quad (12)
\end{aligned}$$

Finally, if $\overline{k} = D$,

$$\mathbf{P}(Q = \overline{k}) = \frac{(N-D)!}{(N-n-1)!(n-D)!} \int_{-\infty}^{\infty} F^D(s) G^{n-D}(s)[1 - G(s)]^{N-n-1} g(s)\, ds, \quad (13)$$

and if $\overline{k} = n$,

$$\mathbf{P}(Q = \overline{k}) = \frac{D!}{(n-1)!(D-n)!} \int_{-\infty}^{\infty} F^{n-1}(s)[1 - F(s)]^{D-n} f(s)[1 - G(s)]^{N-D}\, ds. \quad (14)$$

A derivation of each of these expressions can be found in Appendix A.

5 Robustness with Respect to F and G

5.1 The Standard Hypergeometric Distribution

When $F = G$, Q has a standard hypergeometric distribution. To determine Q we order the random variables from the set $\mathcal{Z}$, sample the n smallest random variables and count the number of random variables from the set $\mathcal{X}$ in the sample. When $F = G$ this is equivalent to sampling n balls from an urn of N balls containing D marked balls without bias. This combinatorial argument leads us to the following theorem.

Theorem 5.1. *When $F = G$, Q has a standard hypergeometric distribution. In particular,*

$$Q \stackrel{d}{=} \mathrm{Hg}(n, D, N). \tag{15}$$

This theorem may alternatively be proved by substituting $F = G$ into the probability mass function derived in Section 4. The details can be found in Appendix B.

What is noteworthy about Theorem 5.1 is that Q is independent of the underlying distributions F and G, when $F = G$.

5.2 Wallenius' Noncentral Hypergeometric Distribution

We also obtain Wallenius' noncentral hypergeometric distribution as a special case.

Theorem 5.2. *When $G = F^{\omega}$, for some $\omega \in (0, \infty)$, Q has Wallenius' noncentral hypergeometric distribution. In particular,*

$$Q \stackrel{d}{=} \mathrm{HgW}(n, D, N, \omega). \tag{16}$$

A proof of this result is given in Appendix C. Notice when $G = F^{\omega}$, the resulting distribution is dependent on F and G only through a single parameter ω. This is a generalisation of the case $F = G$ with $\omega = 1$.

As was the case with Theorem 5.1, the result presented in Theorem 5.2 can be explained intuitively in the context of an urn problem. Recall Q is the number of random variables from the set $\mathcal{X}$ contained in the n smallest random variables from the set $\mathcal{Z}$. When $G = F^{\omega}$, we have

$$\mathbf{P}(X_j < Y_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{y} \omega f(x) F(y)^{\omega-1} f(y)\, dx\, dy = \int_{-\infty}^{\infty} \omega F(y)^{\omega} f(y)\, dy = \frac{\omega}{\omega+1} \tag{17}$$

and

$$\mathbf{P}(Y_i < X_i) = 1 - \frac{\omega}{\omega+1} = \frac{1}{\omega+1}. \tag{18}$$

Thus, this is equivalent to a Wallenius scheme in which n marked balls are sampled from an urn containing D marked balls of weight ω and $N - D$ unmarked balls of weight 1.

6 A General Noncentral Hypergeometric Distribution

From the results presented in Theorem 5.1 and Theorem 5.2, one might anticipate our general noncentral hypergeometric distribution does not depend directly on both F and G. We show in the following theorem that, in fact, it depends only on $F \circ G^{(-1)}$.

Theorem 6.1. *The distribution of Q is parameterised by n, D, N and $F \circ G^{(-1)}$.*

A proof of this theorem can be found in Appendix D. Theorem 6.1 tells us any choices of F and G that give rise to the same function $F \circ G^{(-1)}$ will produce the same general noncentral hypergeometric distribution. This leads us to the following definition.

Definition 6.1. *Suppose a random variable X has the same distribution as the random variable Q. Then we say that X has the* general noncentral hypergeometric distribution *and write $X \stackrel{d}{=} HgG(n, D, N, F \circ G^{(-1)})$, where HgG stands for general hypergeometric.*

We now provide an explicit example in which our noncentral hypergeometric distribution has a distribution other than the standard hypergeometric distribution and Wallenius' noncentral hypergeometric distribution.

Example 6.1. *Consider $X \stackrel{d}{=} HgG(n, D, N, F \circ G^{(-1)})$, where $F(x) = 1 - e^{-\lambda x}$ with $\lambda > 0$ for $x \in [0, \infty)$ and $G(x) = x$ for $x \in [0, 1]$. Using the results from Appendix D, for $\underline{k} < k < \overline{k}$ we have*

$$\mathbf{P}(X = k) = I_1 + I_2, \tag{19}$$

where

$$I_1 = C \int_0^1 t^k (1-t)^{D-k-1} \log^{n-k}(1-t)^{-1/\lambda} [1 - \log(1-t)^{-1/\lambda}]^{N-D-n+k} \, dt \tag{20}$$

and

$$I_2 = \frac{C(N-D-n+k)}{(D-k)} \int_0^1 (1-e^{-\lambda t})^k e^{-\lambda t(D-k)} t^{n-k} (1-t)^{N-D-n+k-1} \, dt, \tag{21}$$

with

$$C = \frac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!}. \tag{22}$$

It is clear X does not have a standard hypergeometric distribution. Applying the binomial theorem to (60) *in Appendix C, shows $\mathbf{P}(X = k)$ may be written as a sum of Beta functions if X has Wallenius' noncentral hypergeometric distribution. If we apply the Binomial theorem*

λ	N	D	n	$\mathbf{E}[X]$	Var(X)	λ	N	D	n	$\mathbf{E}[X]$	Var(X)
0.5	10	5	5	1.39003	0.70331	2	10	5	5	2.84318	0.69158
0.5	10	5	8	3.05486	0.05707	2	10	5	8	3.97031	0.48762
0.5	10	8	5	3.20320	0.19117	2	10	8	5	4.22591	0.42725
0.5	20	10	10	2.87231	1.34529	2	20	10	10	5.70980	1.30273
0.5	20	10	8	2.35258	1.20334	2	20	10	8	4.74683	1.23390
0.5	20	8	10	2.14692	1.18322	2	20	8	10	4.67029	1.24807

Table 1: The mean and variance of X for a variety of parameter values, where F and G are given in Example 6.1.

to (21)*, it may be written as a sum of Kummer's confluent hypergeometric functions (refer to Abramowitz and Stegun (1972)), while* (20) *cannot be written nicely as a sum of well-known functions. Thus, X does not have Wallenius' noncentral hypergeometric distribution either.*

Aside from complete enumeration of all probabilities, there are no exact formulas for the mean and variance of Wallenius' noncentral hypergeometric distribution (Fog, 2008). The general noncentral hypergeometric distribution shares this property. However, given a set of parameter values, the mean and variance may be numerically computed using the probability mass function given in Section 4. Refer to Table 1 for some example Mathematica calculations.

7 Economic Applications

In this section three economic applications for our general noncentral hypergeometric distribution are discussed.

7.1 The $\langle N,M \rangle$ Setup with Unit Traders

We consider a canonical economic model for a market in which a homogeneous, indivisible good is traded among N buyers and M sellers. Each buyer demands at most one unit of the good and sellers have the capacity to produce at most one unit of the good. Assume buyer valuations $V_1, \dots, V_N$ are independent and identically distributed absolutely continuous random variables drawn from the distribution B. Assume seller costs $C_1, \dots, C_M$ are independent and identically distributed absolutely continuous random variables drawn from the distribution S. Furthermore, assume buyer valuations and sellers costs are independent of one another. We let $L = \min\{N, M\}$ and refer to buyers and sellers as agents. An agent's type is its valuation V_i or its cost C_j. The assumption of independently distributed types is standard (but not without loss of generality; see Myerson (1981), Crémer and McLean (1985 and 1988) and McAfee and Reny (1992)) in Bayesian mechanism design.

This unitary $\langle N, M\rangle$ setup has been studied extensively in the literature on Bayesian mechanism design, for example, by Gresik and Satterthwaite (1989), Williams (1999) and Muir (2013). Introductions to mechanism design can be found in Chapter 5 of Krishna's (2010) textbook or in Börgers (2015).

In this setup, we order sellers from lowest cost to highest cost. The lowest cost sellers are more efficient and should trade more often in the market than higher cost sellers. We let $C_{(j)}$ denote the jth most efficient seller.

Buyers are ordered from highest valuation to lowest valuation. Higher valuation buyers are more efficient and should trade more often than lowest valuation buyers. We let $V_{[i]}$ denote the ith most efficient buyer.

7.2 The Walrasian Quantity

Welfare within a market is given by the sum of trading buyer valuations less the the sum of trading seller costs. Market welfare (or social welfare) can be thought of as the total gain from trade experienced by market participants.

A market is said to be efficient if it is such that the welfare of market participants is maximised. The quantity traded in an efficient market is known in economics as the efficient quantity, the welfare-maximsing quantity or the Walrasian quantity. For the $\langle N, M\rangle$ trade setup, we now provide a mathematical definition of the Walrasian quantity.

Definition 7.1. *The* Walrasian quantity*, W, is the quantity traded that maximises social welfare or is efficient. That is,*

$$W = \operatorname*{arg\,max}_{k=0,1,\ldots,L} \sum_{i=1}^{k} (V_{[i]} - C_{(i)}). \tag{23}$$

By virtue of the ordering of buyers and sellers, there is a characterisation of the Walrasian quantity which is more tractable. We have $W = 0$ if and only if $V_{[1]} < C_{(1)}$, and $W = L$ if and only if $V_{[L]} > C_{(L)}$. In all other cases, W is the unique quantity such that

$$V_{[W]} \geq C_{(W)} \quad \text{and} \quad V_{[W+1]} < C_{(W+1)}. \tag{24}$$

We wish to show the Walrasian quantity W is distributed according to our general noncentral hypergeometric distribution.

Lemma 7.1. *For any $i \in \{1, \dots, L\}$, $V_{[i]} > C_{(i)}$ if and only if the N lowest agent types contain at least i seller costs.*

Proof. The result directly follows from applying Lemma 3.1 with appropriate parameter values. $\square$

Lemma 7.2. *W is given by the number of seller costs contained in the N lowest types.*

Proof. If L of the N lowest types are seller costs, clearly $W = L$. Similarly, if none of the N lowest types are seller costs, $W = 0$. Otherwise, W is the unique quantity satisfying (24). By Lemma 7.1, $C_{(W)} \leq V_{[W]}$ implies at least W of the N lowest types are seller costs. Furthermore, $C_{(W+1)} \leq V_{[W+1]}$ implies at least $W + 1$ of the N lowest types are seller costs. Thus, having $C_{[Q]} \leq V_{[W]}$ and $V_{[W+1]} < C_{(W+1)}$ requires that precisely W of the N lowest types are seller costs. $\square$

This leads us to the following theorem.

Theorem 7.1. *W has the general noncentral hypergeometric distribution. Specifically,*

$$W \stackrel{d}{=} HgG(N, M, N + M, S \circ B^{(-1)}). \tag{25}$$

Proof. The result follows directly from Lemma 7.2 and Definition 6.1. $\square$

Theorems 5.1 and 7.1 imply the Walrasian quantity has the standard hypergeometric distribution when $B = S$.

Corollary 7.1. *When $B = S$,*

$$W \stackrel{d}{=} Hg(N, M, N + M). \tag{26}$$

Corollary 7.1 states that, when $B = S$, the distribution of W does not depend on the distribution from which types are drawn. Using mechanism design literature terminology, one would say that W has a prior-free distribution in this case.

7.3 Multi-Unit Traders

A natural extension of the unitary $\langle N, M \rangle$ setup is to consider buyers with multi-unit demand and sellers with multi-unit capacities. Specifically, we now assume each buyer demands ν units of the homogeneous good and each seller has the capacity to produce μ units. Since the good is indivisible, ν and μ must be a positive integers. Assume buyer valuations $V_i^1, \ldots, V_i^\nu$ for each unit of the good are independent and identically distributed with distribution B. Similarly, assume seller costs $C_j^1 \ldots, C_j^\mu$ for each unit of the good are independent and identically distributed with distribution S. If buyers are allowed to trade up to ν units and sellers are allowed to trade up to μ units, this multi-unit setup can simply be considered as a unitary $\langle \nu N, \mu M \rangle$ setup. This multi-unit setup is studied, for example, by Loertscher and Mezzetti (2015). By Theorem 7.1,

$$W \stackrel{d}{=} HgG(\nu N, \mu M, \nu N + \mu M, S \circ B^{(-1)}). \tag{27}$$

Assume next that sellers have unit capacities, which are independent and identically distributed with distribution S. Buyers have multi-unit demands as just described but are restricted to buy one unit at most (for example, because of government imposed rationing). In this case, a buyer's valuation for one unit of the good will be given by

$$V_i = \max\{V_i^1, \ldots, V_i^\nu\}. \tag{28}$$

Thus, we may consider this situation as an $\langle N, M \rangle$ setup in which the distribution of buyer valuations is given by B^ν. Finally, if $B = S$, we obtain a prior-free distribution for the Walrasian quantity.

Corollary 7.2. *Suppose the distribution of buyer valuations is given by B^ν and $S = B$, then*

$$W \stackrel{d}{=} HgW(N, M, N + M, \nu). \tag{29}$$

Proof. The result follows directly from Theorems 5.2 and 7.1. $\square$

7.4 The α-Quantity Traded

Many mechanism design papers (for example Myerson and Satterthwaite (1983)) consider markets in which a convex combination of market welfare, T, and market revenue, R, is maximised. That is, they are designed to maximise $(1 - \alpha)T + \alpha R$, where $\alpha \in [0, 1]$ is a parameter. The quantity traded in such markets, denoted by W_α, is known as the α-quantity traded. Precisely defining and solving this problem in the unitary $\langle N, M\rangle$ setup requires introducing more mechanism design technology than what is necessary for this paper. However, it turns out the α-quantity also has the general noncentral hypergeometric distribution. Start by defining, for $v \in [\underline{v}, \overline{v}]$ and $c \in [\underline{c}, \overline{c}]$, the virtual valuation and cost functions

$$\Phi(v) = v - \frac{1 - B(v)}{b(v)} \quad \text{and} \quad \Gamma(c) = c + \frac{S(c)}{s(c)} \tag{30}$$

and impose the condition that these functions are well-defined and increasing (this is known as Myerson's (1981) regularity condition). Also, for $i \in \{1, \ldots, N\}$ and $j \in \{1, \ldots, M\}$, define weighted virtual valuations and costs

$$V_i^\alpha = (1 - \alpha)V_i + \alpha\Phi(V_i) \quad \text{and} \quad C_j^\alpha = (1 - \alpha)C_j + \alpha\Gamma(C_j). \tag{31}$$

Then it can be shown that the α-quantity traded is equivalent to the Walrasian quantity in the $\langle N, M\rangle$ setup when buyers have valuations given by $V_1^\alpha, \ldots, V_N^\alpha$ and sellers have costs given by $C_1^\alpha, \ldots, C_M^\alpha$. The weighted virtual valuations and costs can be computed from the actual valuations and costs using the weighted virtual type functions,

$$\Phi_\alpha(v) = v - \alpha\frac{1 - B(v)}{b(v)} \quad \text{and} \quad \Gamma_\alpha(c) = c + \alpha\frac{S(c)}{s(c)}. \tag{32}$$

Notice that these functions are well-defined and increasing for $\alpha \in [0, 1]$. Thus, the α-quantity traded has our general noncentral hypergeometric distribution. Specifically,

$$W_\alpha \overset{d}{=} HgG(N, M, N + M, (S \circ \Gamma_\alpha^{(-1)}) \circ (B \circ \Phi_\alpha^{(-1)})^{(-1)}). \tag{33}$$

8 Conclusion

In this paper we constructed a general noncentral hypergeometric distribution from the combined order statistics of two independent and identically distributed samples of random variables with absolutely continuous distribution functions F and G. Our noncentral hypergeometric distributions model biased sampling without replacement.

It was shown that the standard hypergeometric distribution arises as a special case when $F = G$. We obtain Wallenius' noncentral hypergeometric distribution with weight parameter ω when $G = F^\omega$. These results are robust in the sense that we obtain distributions independent of F and G. This motivated us to prove any choices of F and G which give rise to the same function $F \circ G^{(-1)}$ produce the same noncentral hypergeometric distribution.

Thus, our noncentral hypergeometric distribution depends on F and G only through $F \circ G^{(-1)}$.

Finally, it was shown the general noncentral hypergeometric distribution models the distribution of the efficient quantity traded in economic markets. Specifically, in the unitary $\langle N, M \rangle$ setup from Bayesian mechanism design, the Walrasian quantity traded is distributed according to our noncentral hypergeometric distribution. An extension of this model in which buyers had multi-unit demand provided an application of our result regarding Wallenius' noncentral hypergeometric distribution.

Future work could involve developing efficient numerical techniques for computing the moments of the general noncentral hypergeometric distribution. Furthermore, there is scope to investigate asymptotic properties of the distribution. Our noncentral hypergeometric distribution is closely related to the two-sample problem in statistics, which considers comparing the order statistics of two distributions. There exists a body of literature deriving asymptotic results for this problem (see Kochar (2012)). Moreover, there is also scope to apply our general noncentral hypergeometric distribution to additional economic applications as well as some of the biological applications for Wallenius' and Fisher's noncentral hypergeometric distributions. For some of these biological applications, it may be useful to create a multivariate version of our noncentral hypergeometric distribution, as Chesson (1976) did for Wallenius' noncentral hypergeometric distribution.

Acknowledgements

Peter Taylor and Ellen Muir would like to thank the Australian Research Council for supporting this work through the Laureate Fellowship FL130100039 and through the ARC Centre of Excellence for Mathematical and Statistical Frontiers. Ellen Muir would also like to thank the Elizabeth and Vernon Puzey Foundation for support provided through the Elizabeth and Vernon Puzey Scholarship.

References

Abramowitz, M. and Stegun, I. A. (1972), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York, pp. 258,504–505. Tenth printing with corrections.

Börgers, T. (2015), *An Introduction to the Theory of Mechanism Design*, Oxford University Press, New York.

Chesson, J. (1976), 'A non-central multivariate hypergeometric distribution arising from biased sampling with application to selective predation', *Journal of Applied Probability* **13**(4), 795–797.

Crémer, J. and McLean, R. P. (1985), 'Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent', *Econometrica* **53**(2), 345–361.

Crémer, J. and McLean, R. P. (1988), 'Full extraction of the surplus in Bayesian and dominant strategy auctions', *Econometrica* **56**(6), 1247–1257.

David, H. A. (1970), *Order Statistics*, Wiley, New York, chapter 2.

Fisher, R. A. (1935), 'The logic of inductive inference', *The Journal of the Royal Statistical Society* **98**(1), 39–82.

Fog, A. (2008), 'Calculation methods for Wallenius' noncentral hypergeometric distribution', *Communications in Statistics, Simulation and Computation* **37**(2), 258–273.

Gresik, T. A. and Satterthwaite, M. A. (1989), 'The rate at which a simple market converges to efficiency as the number of traders increases: An asymptotic result for optimal trading mechanisms', *Journal of Economic Theory* **48**(1), 304–332.

Kochar, S. C. (2012), 'Stochastic comparisons of order statistics and spacings: A review', *ISRN Probability and Statistics* . Article ID 839473, doi:10.5402/2012/839473.

Krishna, V. (2010), *Auction Theory*, 2 edn, Academic Press, Burlington, Massachusetts, chapter 5.

Loertscher, S. and Mezzetti, C. (2015), Two-sided VCG and double-clock auctions with endogenous reserve prices. Working paper.
URL: *http://simonloertscher.net/data/downloads/12120/LM-DoubAuc270616.pdf*

McAfee, R. P. and Reny, P. J. (1992), 'Correlated information and mechanism design', *Econometrica* **60**(2), 395–421.

Muir, E. V. (2013), Mechanism Design and the ⟨N,M⟩ Trade Problem, Master's thesis, Department of Mathematics and Statistics, The University of Melbourne.

Myerson, R. B. (1981), 'Optimal auction design', *Mathematics of Operations Research* **6**(1), 58–73.

Myerson, R. B. and Satterthwaite, M. A. (1983), 'Efficient mechanisms for bilateral trading', *Journal of Economic Theory* **29**(2), 265–281.

Wallenius, K. T. (1963), Biased Sampling: The Non-central Hypergeometric Probability Distribution, PhD thesis, Department of Statistics, Stanford University. Also published with the same title as Technical Report no. 70.

Williams, S. R. (1999), 'A characterization of efficient, Bayesian incentive compartible mechanisms', *Economic Theory* **14**(1), 155–180.

Appendix A: Derivation of the Probability Mass Function of Q

Order Statistic Results

We require several well-known order statistics results, all of which can be found in David (1970).

First, for $i = 1, \ldots, D$, the density function of the order statistic $X_{(i)}$ is given by

$$f_{X_{(i)}}(x) = \frac{D!}{(i-1)!(D-i)!} F^{i-1}(x)[1-F(x)]^{D-i} f(x), \tag{34}$$

where $x \in \mathbb{R}$. By symmetry, for $j = 1, \ldots, N-D$, the density of the order statistic $Y_{[j]}$ is given by

$$g_{Y_{[j]}}(y) = \frac{(N-D)!}{(j-1)!(N-D-j)!} G^{N-D-j}(y)[1-G(y)]^{j-1} g(y), \tag{35}$$

where $y \in \mathbb{R}$.

Next, for $i = 1, \ldots, D-1$, the joint density function of $(X_{(i)}, X_{(i+1)})$ is given by

$$\begin{aligned} &f_{(X_{(i)},X_{(i+1)})}(x, x') \\ &\quad = \frac{D!}{(i-1)!(D-i-1)!} F^{i-1}(x)[1-F(x')]^{D-i-1} f(x) f(x'), \end{aligned} \tag{36}$$

where $x, x' \in \mathbb{R}$ and $x < x'$. By symmetry, for $j = 1, \ldots, N-D-1$, the joint density function of $(Y_{[j]}, Y_{[j+1]})$ is given by

$$\begin{aligned} &g_{(Y_{[j]},Y_{[j+1]})}(y, y') \\ &\quad = \frac{(N-D)!}{(j-1)!(N-D-j-1)!} G^{N-D-j-1}(y')[1-G(y)]^{j-1} g(y) g(y'), \end{aligned} \tag{37}$$

where $y, y' \in \mathbb{R}$ and $y' < y$.

Proof of (10), (11), (12), (13) and (14)

We start by considering the case $\underline{k} = 0$. Using Theorem 3.1, (34) and (35) we have

$$\begin{aligned} \mathbf{P}(Q = \underline{k}) &= \mathbf{P}(X_{(1)} > Y_{[N-D-n+1]}) \\ &= \int_{-\infty}^{\infty} \int_{y}^{\infty} f_{X_{(1)}}(x) g_{Y_{[N-D-n+1]}}(y)\, dx\, dy \\ &= \int_{-\infty}^{\infty} \int_{y}^{\infty} D[1-F(x)]^{D-1} f(x) g_{Y_{[N-D-n+1]}}(y)\, dx\, dy \\ &= \int_{-\infty}^{\infty} [1-F(y)]^{D} g_{Y_{[N-D-n+1]}}(y)\, dy \\ &= \frac{(N-D)!}{(N-D-n)!(n-1)!} \int_{-\infty}^{\infty} [1-F(s)]^{D} G^{n-1}(s)[1-G(s)]^{N-D-n} g(s)\, ds. \end{aligned}$$

Similarly, when $\underline{k} = n + D - N$,

$$\begin{aligned}
\mathbf{P}(Q = \underline{k}) &= \mathbf{P}(X_{(n+D-N+1)} > Y_{[1]}) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{X_{(n+D-N+1)}}(x) g_{Y_{[1]}}(y)\,dy\,dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{X_{(n+D-N+1)}}(x)(N-D)G^{N-D-1}(y)g(y)\,dy\,dx \\
&= \int_{-\infty}^{\infty} f_{X_{(n+D-N+1)}}(x) G^{N-D}(x)\,dx \\
&= \frac{D!}{(n+D-N)!(N-n-1)!}\int_{-\infty}^{\infty} F^{n+D-N}(s)[1-F(s)]^{N-n-1} f(s) G^{N-D}(s)\,ds.
\end{aligned}$$

We next consider the case $\overline{k} = D$. Using Theorem 3.1, (34) and (35) we obtain

$$\begin{aligned}
\mathbf{P}(Q = \overline{k}) &= \mathbf{P}(Y_{[N-n]} > X_{(D)}) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{y} f_{X_{(D)}}(x) g_{Y_{[N-n]}}(y)\,dx\,dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{y} DF^{D-1}(x) f(x) g_{Y_{[N-n]}}(y)\,dx\,dy \\
&= \int_{-\infty}^{\infty} F^{D}(y) g_{Y_{[N-n]}}(y)\,dy \\
&= \frac{(N-D)!}{(N-n-1)!(n-D)!}\int_{-\infty}^{\infty} F^{D}(s) G^{n-D}(s)[1-G(s)]^{N-n-1} g(s)\,ds.
\end{aligned}$$

Similarly, when $\overline{k} = n$,

$$\begin{aligned}
\mathbf{P}(Q = \overline{k}) &= \mathbf{P}(Y_{[N-D]} > X_{(n)}) \\
&= \int_{-\infty}^{\infty}\int_{x}^{\infty} f_{X_{(n)}}(x) g_{Y_{[N-D]}}(y)\,dy\,dx \\
&= \int_{-\infty}^{\infty}\int_{x}^{\infty} f_{X_{(n)}}(x)(N-D)[1-G(y)]^{N-D-1} g(y)\,dy\,dx \\
&= \int_{-\infty}^{\infty} f_{X_{(n)}}(x)[1-G(x)]^{N-D}\,dx \\
&= \frac{D!}{(n-1)!(D-n)!}\int_{-\infty}^{\infty} F^{n-1}(s)[1-F(s)]^{D-n} f(s)[1-G(s)]^{N-D}\,ds.
\end{aligned}$$

Finally, take $\underline{k} < k < \overline{k}$ and by Theorem 3.1 we have

$$\begin{aligned}
\mathbf{P}(Q = k) &= \mathbf{P}(Y_{[N-D-n+k]} > X_{(k)}, X_{(k+1)} > Y_{[N-D-n+k+1]}) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{x'}\int_{y'}^{\infty}\int_{-\infty}^{\min\{y,x'\}} f_{(X_{(k)},X_{(k+1)})}(x,x')g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dx\,dy\,dy'\,dx' \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{x'}\int_{\max\{y',x'\}}^{\infty}\int_{-\infty}^{x'} f_{(X_{(k)},X_{(k+1)})}(x,x')g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dx\,dy\,dy'\,dx' \\
&\quad + \int_{-\infty}^{\infty}\int_{-\infty}^{x'}\int_{y'}^{x'}\int_{-\infty}^{y} f_{(X_{(k)},X_{(k+1)})}(x,x')g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dx\,dy\,dy'\,dx' \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{x'}\int_{x'}^{\infty}\int_{-\infty}^{x'} f_{(X_{(k)},X_{(k+1)})}(x,x')g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dx\,dy\,dy'\,dx' \\
&\quad + \int_{-\infty}^{\infty}\int_{y}^{\infty}\int_{-\infty}^{y}\int_{-\infty}^{y} f_{(X_{(k)},X_{(k+1)})}(x,x')g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dx\,dy'\,dx'\,dy.
\end{aligned} \tag{38}$$

To simplify notation, let

$$I_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{x'}\int_{x'}^{\infty}\int_{-\infty}^{x'} f_{(X_{(k)},X_{(k+1)})}(x,x')g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dx\,dy\,dy'\,dx' \tag{39}$$

and

$$I_2 = \int_{-\infty}^{\infty}\int_{y}^{\infty}\int_{-\infty}^{y}\int_{-\infty}^{y} f_{(X_{(k)},X_{(k+1)})}(x,x')g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dx\,dy'\,dx'\,dy, \tag{40}$$

so (38) becomes

$$\mathbf{P}(Q = k) = I_1 + I_2. \tag{41}$$

Recomputing I_1 using (36) and (37) gives

$$\begin{aligned}
I_1 &= \tfrac{D!}{(k-1)!(D-k-1)!} \int_{-\infty}^{\infty}\int_{-\infty}^{x'}\int_{x'}^{\infty}\int_{-\infty}^{x'} F^{k-1}(x)[1-F(x')]^{D-k-1} f(x) f(x') g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dx\,dy\,dy'\,dx' \\
&= \tfrac{D!}{k!(D-k-1)!} \int_{-\infty}^{\infty}\int_{-\infty}^{x'}\int_{x'}^{\infty} F^{k}(x')[1-F(x')]^{D-k-1} f(x') g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dy\,dy'\,dx' \\
&= \tfrac{D!(N-D)!}{k!(n-k-1)!(D-k-1)!(N-D-n+k-1)!} \int_{-\infty}^{\infty}\int_{-\infty}^{x'}\int_{x'}^{\infty} F^{k}(x')[1-F(x')]^{D-k-1} f(x') G^{n-k-1}(y')[1-G(y)]^{N-D-n+k-1} g(y)g(y')\,dy\,dy'\,dx' \\
&= \tfrac{D!(N-D)!}{k!(n-k-1)!(D-k-1)!(N-D-n+k)!} \int_{-\infty}^{\infty}\int_{-\infty}^{x'} F^{k}(x')[1-F(x')]^{D-k-1} f(x') G^{n-k-1}(y')[1-G(x')]^{N-D-n+k} g(y')\,dy'\,dx' \\
&= \tfrac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!} \int_{-\infty}^{\infty} F^{k}(x')[1-F(x')]^{D-k-1} f(x') G^{n-k}(x')[1-G(x')]^{N-D-n+k}\,dx'.
\end{aligned} \tag{42}$$

Recomputing I_2 using (36) and (37) gives

$$\begin{aligned}
I_2 &= \tfrac{D!}{(k-1)!(D-k-1)!} \int_{-\infty}^{\infty}\int_{y}^{\infty}\int_{-\infty}^{y}\int_{-\infty}^{y} F^{k-1}(x)[1-F(x')]^{D-k-1} f(x) f(x') g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dx\,dy'\,dx'\,dy \\
&= \tfrac{D!}{k!(D-k-1)!} \int_{-\infty}^{\infty}\int_{y}^{\infty}\int_{-\infty}^{y} F^{k}(y)[1-F(x')]^{D-k-1} f(x') g_{(Y_{[N-D-n+k]},Y_{[N-D-n+k+1]})}(y,y')\,dy'\,dx'\,dy \\
&= \tfrac{D!(N-D)!}{k!(n-k-1)!(D-k-1)!(N-D-n+k-1)!} \int_{-\infty}^{\infty}\int_{y}^{\infty}\int_{-\infty}^{y} F^{k}(y)[1-F(x')]^{D-k-1} f(x') G^{n-k-1}(y')[1-G(y)]^{N-D-n+k-1} g(y)g(y')\,dy'\,dx'\,dy \\
&= \tfrac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k-1)!} \int_{-\infty}^{\infty}\int_{y}^{\infty} F^{k}(y)[1-F(x')]^{D-k-1} f(x') G^{n-k}(y)[1-G(y)]^{N-D-n+k-1} g(y)\,dx'\,dy \\
&= \tfrac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k-1)!} \int_{-\infty}^{\infty} F^{k}(y)[1-F(y)]^{D-k} G^{n-k}(y)[1-G(y)]^{N-D-n+k-1} g(y)\,dy.
\end{aligned} \tag{43}$$

Combining (41), (42), (43) we have

$$\begin{aligned}
\mathbf{P}(Q=k) = \tfrac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!}\Bigg(&\int_{-\infty}^{\infty} F^{k}(s)[1-F(s)]^{D-k-1} f(s) G^{n-k}(s)[1-G(s)]^{N-D-n+k}\,ds \\
&+ \tfrac{N-D-n+k}{D-k} \int_{-\infty}^{\infty} F^{k}(s)[1-F(s)]^{D-k} G^{n-k}(s)[1-G(s)]^{N-D-n+k-1} g(s)\,ds \Bigg).
\end{aligned} \tag{44}$$

Appendix B: Proof of Theorem 5.1

To prove Theorem 5.1, we start with the probability mass function of Q and impose the requirement that $F = G$. The resulting expression is then rewritten in terms the Beta function, denoted by B. Finally, everything is written in terms of factorials using the relationship between the Beta and Gamma function (refer to (Abramowitz and Stegun, 1972)) and simplified.

Proof. Take $\underline{k} < k < \overline{k}$ and compute the integral terms in (12). We write

$$\mathbf{P}(Q = k) = I_1 + I_2, \tag{45}$$

where I_1 and I_2 are defined in (42) and (43). Using $F = G$, I_1 becomes

$$I_1 = \frac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!} \int_{-\infty}^{\infty} F^n(s)[1-F(s)]^{N-n-1} f(s)\, ds. \tag{46}$$

Making the change of variables $t = F(s)$ gives

$$\begin{aligned} I_1 &= \frac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!} \int_0^1 t^n(1-t)^{N-n-1}\, dt \\ &= \frac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!} B(n+1, N-n) \\ &= \frac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!} \frac{n!(N-n-1)!}{N!} \\ &= \frac{D!n!(N-D)!(N-n-1)!}{k!N!(n-k)!(D-k-1)!(N-D-n+k)!}. \end{aligned} \tag{47}$$

Using $F = G$, I_2 becomes

$$I_2 = \frac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k-1)!} \int_{-\infty}^{\infty} G^n(s)[1-G(s)]^{N-n-1} g(s)\, ds. \tag{48}$$

Making the change of variables $t = G(s)$ we obtain

$$\begin{aligned} I_2 &= \frac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k-1)!} \int_0^1 t^n(1-t)^{N-n-1}\, dt \\ &= \frac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k-1)!} B(n+1, N-n) \\ &= \frac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k-1)!} \frac{n!(N-n-1)!}{N!} \\ &= \frac{D!n!(N-D)!(N-n-1)!}{k!N!(n-k)!(D-k)!(N-D-n+k-1)!}. \end{aligned} \tag{49}$$

Finally, substituting (47) and (49) into (45) gives

$$\begin{aligned}\mathbf{P}(Q=k) &= \frac{D!n!(N-D)!(N-n-1)!\,[(D-k)+(N-D-n+k)]}{k!N!(n-k)!(D-k)!(N-D-n+k)!} \\ &= \frac{D!n!(N-D)!(N-n)!}{k!N!(n-k)!(D-k)!(N-D-n+k)!} \\ &= \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}}. \end{aligned} \tag{50}$$

The $Q=\underline{k}$ and $Q=\overline{k}$ cases are similar. Thus, Q has the standard hypergeometric distribution when $F=G$ and

$$Q \stackrel{d}{=} \mathrm{Hg}(n,D,N). \tag{51}$$

□

Appendix C: Proof of Theorem 5.2

Suppose $R \stackrel{d}{=} HgW(n,D,N,\omega)$. From (20) in Fog (2008) the probability mass function of R may be written

$$\mathbf{P}(R=r) = \binom{D}{r}\binom{N-D}{n-r} d\int_0^1 (1-t^{\omega})^r(1-t)^{n-r}t^{d-1}\,dt, \tag{52}$$

where

$$d = N-D-n+r+\omega(n-r) \tag{53}$$

and r lies in the range $\{\max\{0,n+D-N\},\ldots,\min\{n,D\}\}$.

In the context of an urn problem, the number of marked balls included in a sample also characterises the number of marked balls not included. The probability of sampling r marked balls of weight ω in n draws is equal to the probability of sampling $D-r$ marked balls of weight $1/\omega$ in $N-n$ draws. Thus, the density function for R may be rewritten

$$\mathbf{P}(R=r) = \binom{D}{r}\binom{N-D}{n-r} d\int_0^1 (1-t^{1/\omega})^{D-r}(1-t)^{N-D-n+r}t^{d-1}\,dt \tag{54}$$

where

$$d = n-r+r/\omega \tag{55}$$

and r lies in the range $\max\{0,n+D-N\},\ldots,\min\{n,D\}$.

We prove when $G=F^{\omega}$, Q has the same distribution as the random variable R.

Proof. Take $\underline{k} < k < \overline{k}$ and compute $\mathbf{P}(Q = k)$. We write

$$\mathbf{P}(Q = k) = I_1 + I_2, \tag{56}$$

where I_1 and I_2 are defined in (42) and (43).

Making the change of variables $t = G(s) = F^\omega(s)$, (42) becomes

$$\begin{aligned} I_1 &= \tfrac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!}\tfrac{1}{\omega}\int_0^1 t^{k/\omega}(1-t^{1/\omega})^{D-k-1}t^{1/\omega-1}t^{n-k}(1-t)^{N-D-n+k}\,dt \\ &= \binom{D}{k}\binom{N-D}{n-k}(D-k)/\omega\int_0^1 (1-t^{1/\omega})^{D-k-1}(1-t)^{N-D-n+k}t^{n-k+(k+1)/\omega-1}\,dt. \end{aligned} \tag{57}$$

Similarly, making the change of variables $t = G(s) = F^\omega(s)$ in (43) gives

$$\begin{aligned} I_2 &= \tfrac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k-1)!}\int_0^1 t^{k/\omega}(1-t^{1/\omega})^{D-k}t^{n-k}(1-t)^{N-D-n+k-1}\,dt \\ &= \tfrac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k-1)!}\int_0^1 (1-t^{1/\omega})^{D-k}(1-t)^{N-D-n+k-1}t^{n-k+k/\omega}\,dt \\ &= \tfrac{-D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k)!}\left[(1-t^{1/\omega})^{D-k}(1-t)^{N-D-n+k}t^{n-k+k/\omega}\right]_0^1 \\ &\quad + \tfrac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k)!}(n-k+k/\omega)\int_0^1 (1-t^{1/\omega})^{D-k}(1-t)^{N-D-n+k}t^{n-k+k/\omega-1}\,dt \\ &\quad - \tfrac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k)!}(D-k)/\omega\int_0^1 (1-t^{1/\omega})^{D-k-1}(1-t)^{N-D-n+k}t^{n-k+(k+1)/\omega-1}\,dt \\ &= \binom{D}{k}\binom{N-D}{n-k}(n-k+k/\omega)\int_0^1 (1-t^{1/\omega})^{D-k}(1-t)^{N-D-n+k}t^{n-k+k/\omega-1}\,dt \\ &\quad - \binom{D}{k}\binom{N-D}{n-k}(D-k)/\omega\int_0^1 (1-t^{1/\omega})^{D-k-1}(1-t)^{N-D-n+k}t^{n-k+(k+1)/\omega-1}\,dt. \end{aligned} \tag{58}$$

Substituting (57) and (58) into (56) gives

$$\begin{aligned} &\mathbf{P}(Q = k) \\ &\quad = \binom{D}{k}\binom{N-D}{n-k}(n-k+k/\omega)\int_0^1 (1-t^{1/\omega})^{D-k}(1-t)^{N-D-n+k}t^{n-k+k/\omega-1}\,dt. \end{aligned} \tag{59}$$

When $r = k$ in (55), $d = n - k + k/\omega$, and so

$$\mathbf{P}(Q = k) = \binom{D}{k}\binom{N-D}{n-k}d\int_0^1 (1-t^{1/\omega})^{D-k}(1-t)^{N-D-n+k}t^{d-1}\,dt, \tag{60}$$

which is identical to (54) when $r = k$ as required. The $Q = \underline{k}$ and $Q = \overline{k}$ cases are similar. □

Appendix D: Proof of Theorem 6.1

We show for general F and G, the resulting noncentral hypergeometric distribution is parameterised by n, D, N and $F \circ G^{(-1)}$.

Proof. For $\underline{k} < k < \overline{k}$, we have

$$\mathbf{P}(Q = k) = I_1 + I_2, \tag{61}$$

where I_1 and I_2 are defined in (42) and (43). Starting from (42) and making the change of variables $t = F(s)$ gives

$$\begin{aligned} I_1 &= C_1 \int_{-\infty}^{\infty} F^k(s)[1 - F(s)]^{D-k-1} f(s)[G(F^{(-1)}(F(s)))]^{n-k}[1 - G(F^{(-1)}(F(s)))]^{N-D-n+k}\, ds \\ &= C_1 \int_0^1 t^k (1-t)^{D-k-1}[G(F^{(-1)}(t))]^{n-k}[1 - G(F^{(-1)}(t))]^{N-D-n+k}\, dt \\ &= C_1 \int_0^1 t^k (1-t)^{D-k-1}[(F \circ G^{(-1)})^{(-1)}(t)]^{n-k}[1 - (F \circ G^{(-1)})^{(-1)}(t)]^{N-D-n+k}\, dt, \end{aligned} \tag{62}$$

where

$$C_1 = \frac{D!(N-D)!}{k!(n-k)!(D-k-1)!(N-D-n+k)!}. \tag{63}$$

Similarly, making the change of variables $t = G(s)$ in (43) we have

$$\begin{aligned} I_2 &= C_2 \int_{-\infty}^{\infty} [F(G^{(-1)}(G(s)))]^k[1 - F(G^{(-1)}(G(s)))]^{D-k} G^{n-k}(s)[1 - G(s)]^{N-D-n+k-1} g(s)\, ds \\ &= C_2 \int_0^1 [F(G^{(-1)}(t))]^k[1 - F(G^{(-1)}(t))]^{D-k} t^{n-k}(1-t)^{N-D-n+k-1}\, dt \\ &= C_2 \int_0^1 [(F \circ G^{(-1)})(t)]^k[1 - (F \circ G^{(-1)})(t)]^{D-k} t^{n-k}(1-t)^{N-D-n+k-1}\, dt, \end{aligned} \tag{64}$$

where

$$C_2 = \frac{D!(N-D)!}{k!(n-k)!(D-k)!(N-D-n+k-1)!}. \tag{65}$$

From (62) and (64) it can be seen that in this case the distribution of Q depends only on n, D, N and $F \circ G^{(-1)}$ as required. The $Q = \underline{k}$ and $Q = \overline{k}$ cases are similar. □