ON THE DISTRIBUTION OF THE MAXIMUM OF n INDEPENDENT NORMAL RANDOM VARIABLES: IID AND INID CASES

[1]Roberto N. Padua, [2]Mark S. Borres and [3]Razil M. Gumanoy

ABSTRACT

The paper deals with the distribution of the maximum of n independent normal random variables and hints on some of its applications in the electricity power industry in the area of peak load estimation and in genetic selection for animal breeding. The paper provides for simple approximations to the mean of the largest order statistics both in the iid and non-identically distributed cases. Likewise, while the large sample results for the iid case have been treated in the past, we focused on the relatively unexplored non-identical but independent case. Large sample asymptotic results for extreme values of normal random variables are often used in reliability theory and also used in the analysis of extreme weather changes in relation to climate change. Results show that the large sample distribution for non-identically distributed case still obeys the Type I Gumbel distribution with shifted parameters through an application of Frechet's stability postulate.

Keywords: *largest order statistic, multivariate normal, error function , peak load, Rayleigh Distribution, Gumbel Distribution*

1.0 Introduction

The hourly load demand for electricity as noted by an Electric Cooperative roughly follows a normal distribution so that if X_j represents the demand at hour j, then $X_j \sim N(\mu_j, \sigma_j^2)$. We seek the probability that the peak demand occurs at time t, that is, we want to evaluate:

$$1 \ldots \; Pr(X_t > Y),$$

where Y is the maximum of the X's. Such a practical problem occurs almost daily in most industries the country that the need to develop analytic methods to tackle it is almost imperative. In the case that there are only two (2) potential peak hours, X_1 and X_2, then the probability that X_1 is greater than X_2 can be easily calculated. In fact, this is equivalent to finding the probability, $Pr(X_1 - X_2 > 0)$ which can be obtained from the distribution $N(\mu_1-\mu_2, \sigma_1^2+\sigma_2^2)$ by a table look up or actual numerical integration. When there are n competing hours, then one must resolve Equation (1) which is the focus of the present paper.

The problem of finding the distribution of Y is, by itself, nothing new. In fact, the classical approach would be to consider:

$$(2) \ldots Pr(Y \le y) = Pr(X_1 \le y, X_2 \le y, \ \ldots X_n \le y) = \textstyle\prod_i Pr(X_i \le y) = \prod_i \Phi_i(y)$$

where $\Phi_i(.)$ is the cumulative distribution

[1]Liceo de Cagayan University
[2]University of San Jose Recoletos
[3]Surigao del Sur State University

function of a normal distribution with mean μ_i and variance σ_i^2. The probability problem given by (1) reduces to:

$$3\ldots\ Pr(X_t \leq Y) = 1 - Pr(X_t \leq Y|Y) \prod_y Pr(Y \leq y).$$

Equation (3) demands a large amount of integration and we shall not approach the problem this way. A related problem arises out of this basic problem of determining the probability distribution of the maximum of n independent normal random variables. Consider the daily peak loads noted $Y_1, Y_2, \ldots, Y_n$ for a period n. The distribution of each Y_i is given by (2) and will be more explicitly stated in the body of this paper. We wish to know the joint distribution of these peak loads $f(Y_1, Y_2, .., Y_n)$. Knowledge of the joint distribution $f(.)$ allows us to ask relevant questions such as, what is the probability that the daily peak loads do not exceed a capacity limit L? That is, we seek answer to the probability question:

$$4\ldots.\ P_y\,(Y_1 \leq L,\ Y_2 \leq L, \ldots Y_n \leq L)?$$

With the passage of the Electricity Power Industry Reform Act (1998), the issues of generation capacity and demand requirements, average and peak load requirements and others had been highlighted because of the unbundling of the electricity charges passed on to the consumers. For instance, distribution utilities have to contend with the issue of determining how much electricity to procure from power generation companies considering both the average demand and the peak load demand in their service areas. A miscalculation on the part of the distribution utilities could mean millions in terms of losses.

Hill (2010) considered a similar problem in relation to the calculation of the probability that a runner wins in an n-player running match. Specifically, he obtained the probability distribution of the minimum of n independent normal random variable. He found that the probability distribution of Y, the minimum of n independent normal random variables obeys a multivariate normal distribution with mean vector μ and covariance matrix S where:

$$5\ldots\ \mu = (\mu_1, \mu_2, \ldots, \mu_n)' \text{ and } S = \text{diag}(\sigma_i^2)$$

is a diagonal matric with σ_i^2 on the i^{th} diagonal and zeroes elsewhere, or

$$6\ldots\ f(t) = \frac{1}{\sqrt{\det S}(\sqrt{2\pi})^n} e^{-\frac{1}{2}(t-\mu)S^{-1}(t-\mu)^T}$$

The application of Hill's (2010) results in more practical areas such as the power industry sector is obvious. One might be interested in the downtime power load (instead of the peak power load) for purposes of planning and forecasting of a distribution utility's daily demand requirements.

Finally, a related problem that might be of interest is the probability distribution of the maximum of the maxima of random variables. Let

$S_1 = \{x_{11}, x_{12}, \ldots, x_{1n}\}$,
$S_2 = \{x_{21}, x_{22}, \ldots, x_{2n}\}, \ldots,$
$S_p = \{x_{1p}, x_{2p}, \ldots, x_{pn}\}$

be subsets of independent random variables of equal length n. Let :

$$7\ldots\ Y_i = \max_{x \in S_i}\{x_{ij}\}$$
$$T = \max\{Y_i\},\ i = 1, 2, , \ldots, p$$

2.0 Mathematical Derivation of the Distribution of the Maxima of n Normal Random Variables

We are interested in evaluating P_Y (X_0

> Y) where Y = max $\{X_j\}$ j = 1,2,..,n with the same assumptions as Hill's (2010) paper. Following his derivation, we first obtain the probability distribution of Y:

$$(8)\quad \Pr(Y \le a) = \prod_j P(X_j \le a)$$
$$= \prod_j \left(\frac{1}{2} + erf\left(\frac{a-\mu_j}{\sigma_j\sqrt{2}}\right)\right)$$
$$= \int_{-\infty}^{\frac{a-\mu_1}{\sigma_1\sqrt{2}}} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}x_1^2} \ldots \int \int_{-\infty}^{\frac{a-\mu_n}{\sigma_n\sqrt{2}}} \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{1}{2}x_n^2}\, dx_1 \ldots dx_n$$
$$= \int f(t)dt = \int \frac{1}{\sqrt{\det(S)}(\sqrt{2\pi})^n} e^{-\frac{1}{2}(t-\mu)' S^{-1}(t-\mu)} dt$$

Hence, the probability distribution of the maximum of n independent normal random variables is precisely a multivariate normal distribution where:

$$t = (X_1, X_2, \ldots, X_n)', \qquad S = diag(\sigma_i^2), \qquad \mu = (\mu_1, \mu_2, \ldots, \mu_n).$$

Let us write down the density function of X_0 g (x) as:

$$\frac{1}{\sigma_0\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu 0)^2/\sigma_0^2}$$

The desired probability can now be obtained using (8):

$$(9)\quad Pr(X_0 > Y) = Pr(Y < X_0) = \int P(Y < X)P(X = x_0)\, dx = \int g(x_0)dx_0 \int f(t)\, dt$$
$$= \iint_Y^{\infty} g(x_0) f(t)\, dx_0\, dt \text{ and the outer integral is from } (-\infty, \infty).$$

However, the integral is precisely the integral of an (n+1)-variate normal distribution or the n-variate normal distribution before plus one additional variable. That is, h(t) is:

$$10\ldots \frac{1}{\sqrt{\det(S)}(\sqrt{2\pi})^{n+1}} e^{-\frac{1}{2}(t-\mu)' S^{-1}(t-\mu)}$$

where S is an (n+1) x (n+1) strictly diagonal positive definite matrix whose diagonal elements are σ_i^2, i= 0,1,...,n and $\mu = (\mu_0, \mu_1, \ldots, \mu_n)$. The computation of the integral , however, is not a trivial task and we shall return to this issue later when we perform our simulation exercises.

Application 1: Given the hourly data for electricity demand in a given locality, we can ask the question of finding the probability that the peak demand is less than 50MW. If $X_1, X_2, .., X_{24}$ are the hourly electricity demands, and Y = max $\{X_i\}$, then we can compute P(Y ≤ 50) using Equation (8). Of course, we need the hourly data (24-hour data) over at least one month to establish the distribution of the X's.

Application 2: In genetics, Rawlings (1976), Hill(1976,1977) and Tong(1990), considered the problem of selecting the best animal out of n animals for breeding purposes. Let $X_1; \ldots ; X_n$ be the measurements of a certain biological or physical characteristic of the n animals, such as the body weights or back fats of the pigs. The animal with score $Y = X_{(n)}$ is to be selected. If $X_1; \ldots ; X_n$ are independent with mean μ, then the common mean of the observations of offspring of the selected animal with score $X_{(n)}$ is $E(X_{(n)})$, and therefore the expected gain in one generation is $E(X_{(n)}) - \mu$. The animals must come from different stocks because the assumption of independence will be violated. If Y = max $\{X_i\}$ we can ask the same probability question as before.

Independent and Identically (IID) Normal Random Variables Case

Often, our interest rests mainly on the mean of the maximum of n random variables. Even in the case where the random variables are normally distributed, the numerical integration required to determine the expected value of the maximum is tedious. It is possible to develop heuristic approximations to the expected value of the maximum of n iid normal random variables. Let $X_{(n)}$ = max $\{X_i\}$. Clearly $X_{(n)} > E(X_i) = \mu$ for all i. Suppose that:

11...$X_{(n)} = \mu + \varepsilon$ where $\varepsilon \sim F(\varepsilon)$, and
$F(\varepsilon) = 1 - \exp(-\frac{\varepsilon^2}{2b^2})$, $0 < \varepsilon < \infty$

The distribution *f* (.) is called a Rayleigh Distribution used in extreme value analysis and is a special case of the Weibull distribution with $\alpha = \sqrt{2}b$ and $\beta=2$. Model 11 says that the maximum order statistic $X_{(n)}$ exceeds the mean of the component normal random variables by a random amount ε whose expected value is:

12.... $E(\varepsilon) = 2^{½} b\, \Gamma(1+½)$

The expected value of $X_{(n)}$ can be obtained from (11):

13.... $E(X_{(n)}) = \mu + E(\varepsilon)$.

In other words, if we can model the extreme value statistic $X_{(n)}$ as a sum of the common mean plus a Rayleigh distributed random error, then it is possible to estimate its mean as well by (13).

Procedurally, if $x_1, x_2, .., x_n$ are iid N (μ, σ^2) are random samples, then we first take the maximum likelihood estimators of μ and σ^2 corrected for bias:

14... $\bar{x} = n^{-1}\sum x_i$
$s^2 = (n-1)^{-1}\sum (x_i - \bar{x})^2$

Put $\varepsilon_i = |x_i - \bar{x}|$ for i = 1, 2,...,n, which we now assume obeys the Rayleigh distribution with mean equal to (13). The maximum likelihood estimator of the parameter b of the Rayleigh distribution is given by:

15...$b = \sqrt{\frac{1}{2n}\sum_1^n \epsilon_i^2}$

Our estimate for $E(X_{(n)}) = \mu_{(n)}$ is thus:

(16) $\widehat{\mu_{(n)}} = \bar{x} + 2\sqrt{\pi}\sqrt{\frac{1}{2n}\sum_1^n \epsilon_i^2} = \bar{x} + \sqrt{2\pi}\sqrt{\frac{1}{n}\sum_1^n \epsilon_i^2} \approx \bar{x} + 2.51\, sd(x)$

or heuristically, $\widehat{\mu_{(n)}} = \bar{x} + 3$ sd (x) which is intuitively appealing. We verify this error model by using simulation in the next section.

Independent But Not Identically Distributed Normal Random Variable

The situation when the random variables X_i are not identically distributed but independent normal random variables with $\mu_i = E(X_i)$ and σ_i^2 = var (X_i) is more complicated but also more useful in practice. In the succeeding discussions, we agree on the following notation:

$\bar{\mu} = \frac{1}{n}\sum \mu_i$ = mean of the means
$\mu_{(n)} = \max\{\mu_1, \mu_2, \ldots, \mu_n\}$ = maximum of the means
$X_{(n)} = \max.\{X_1, X_2, \ldots, X_n\}$ = maximum order statistic $D_n = \mu_{(n)} - \bar{\mu}$.

We create a data model similar to (11). We start with the obvious inequality:

17...$\mu_{(n)} - \bar{\mu} \geq 0$ with equality only if $\mu_i = \mu$ for all *i*. To see this, consider:

18...$\mu_{(n)} = \frac{1}{n}\sum_1^n \mu_{(n)} > \frac{1}{n}\sum_1^n \mu_i = \bar{\mu}$

Since $\mu_{(n)} > \mu_i$ for all *i*. It follows that $D_n \geq 0$. We now claim that :

19...$X_{(n)} = \bar{\mu} + \varepsilon_i$ where

$\varepsilon_i = |x_i - \mu_{(n)}|$, i = 1,2,..,n. and : $\varepsilon_i \sim F(\varepsilon)$ is a Rayleigh distribution function. In practice, the underlying means are unknown and so we replace (19) by its sample counterpart:

20.... $X_{(n)} = \bar{x} + |x_i - x_{(n)}|$.

The second term on the right is assumed to obey a Rayleigh distribution with mean given by Equation (12). However, the MLE of b is now:

21.... $\hat{b} = \sqrt{\frac{1}{2n}\sum_1^n |x_i - \mu_{(n)}|^2}$.

Equation (21) is no longer approximately equal to $2^{½}$sd(x). We can argue heuristically to obtain a sense of the magnitude of (21) in relation to the case when the x_i's are properly centered around their means.

Let $Y = (x_1-\mu_1)^2 + (x_2 - \mu_2)^2 + \ldots + (x_n - \mu_n)^2$, and $Z = (x_1-\mu_{(n)})^2 + (x_2 - \mu_{(n)})^2 + \ldots + (x_n - \mu_{(n)})^2$. If we take expectations:

22....$E(Y) = \sigma^2 + \sigma^2 + \ldots + \sigma^2 = n\sigma^2$,

assuming equal variances. Next consider one term of the quantity Z:

$$(23)\ E(x_i - \mu_{(n)})^2 = E(x_i - \mu_i + \mu_i - \mu_{(n)})^2$$
$$= E(x_i-\mu_i)^2 + 2E(x_i - \mu_i)(\mu_i - \mu_{(n)}) + E(\mu_i - \mu_{(n)})^2$$
$$= \sigma^2 + 0 + \theta_i^2,$$

since the second term above is zero and

$$\theta_1^2 = (\mu_i - \mu_{(n)})^2$$

It follows that $E(Y) = n\sigma^2 \leq E(Z) = n\sigma^2 + \sum_1^n \theta_i^2$. Equation (21) is greater than Equation (15), and so we expect a greater additive factor to expected value of the means than in the iid case. In fact, the inequality provides us an insight on the magnitude of the difference since:

24.... $E(Z) - E(Y) = \sum_1^n \theta_i^2$.

There is a rough approximation on the value of $E(X_{(n)})$ provided by Hamza (2008) and we borrow his theorem below:

Theorem (Hamza). Let $X_1, X_2, \ldots, X_n$ be independent random variables with $M_i = E(X_i)$, then:

$$\bar{M} \leq E(X_{(n)}) \leq \bar{M} + \frac{n-1}{n} M_{(n)}.$$

where:

$\bar{M}$ = average of the M_i's

$M_{(n)} = \max\{M_i\}$.

In relation to the present problem, it may often be more useful for the electric distribution utility to have an idea of the magnitude of the peak demand (on the average). Thus, Hamza(s) (2008) theorem will be most useful in providing such information. Similarly, in the genetic selection problem of application 2, we can assume that the means differ across the animals and that we chose the maximum observed X_i as the animal to be used for breeding. Then, again, Hamza's (2008) results will apply. Note that the Theorem does not require that the random variables be normal. It applies to all independent random variables, and so, is quite general. Meanwhile for sufficiently large n, we can consider some asymptotic results to simplify the calculations.

Asymptotic Results for the IID Case

We wish to show that for large n, the asymptotic distribution of the largest order statistic $X_{(n)}$ is a Gumbel Type I distribution. According to the Fisher-Tippet–Gnedenko theorem, the maximum of a sample of iid random variables after proper "whitening" converges in distribution to one of three possible distributions, the Gumbel distribution, the Fréchet distribution, or the Weibull distribution. The role of extremal types theorem for maxima is similar to that of central limit theorem for averages.

Let Y = max $\{X_i\}$ of a sequence of independent and identically distributed standard normal random variables X_i. Let $\Phi(x)$ denote the cumulative distribution function of a standard normal random variable x. The cumulative distribution function of the maximum order statistic Y is:

25...$F_n(y) = [\Phi(y)]^n$, $-\infty < y < \infty$ as before.

Clearly, lim $F_n(y) = 1$ or 0 depending on whether $\Phi(y) = 1$ or 0 as $n \to \infty$. In order to obtain a non-degenerate limiting distribution, it is necessary to transform Y by applying a linear transformation with coefficients which depend on the sample size n but not on y. This process is similar to the standardization process in statistics. Let $Y_n' = a_n Y + b_n$ where a_n and b_n are coefficients depending on n but not on y. Suppose first that the limiting distribution G(y) exists. That is: lim $F_n(y) = G(y)$ exists for properly transformed Y.

If we increase the sample size to nN where N > 0, then the largest of the nN values $X_1, X_2, \ldots, X_{nN}$ is also equal to the largest of the values $X_{j-1(n+1)}, X_{j-1(n+2)}, \ldots\ldots, X_{jn}$, for j=1,2,...,N. It follows that the limiting distribution G(.) obeys:

26... $[G(y)]^N = G(a_N y + b_N)$, Frechet (1927)

Equation (26) is called the *stability postulate*. Now, take $a_N = 1$, Equation (26) now becomes:

27... $[G(y)]^N = G(y + b_N)$.

We iterate Equation (27) for larger samples NM:

28 ... $[G(y)]^{NM} = G(x + b_{NM}) = G(y + b_N)^M = G(y + bN + bM)$.

We infer that : $b_{NM} = b_N + b_M$. This equation tells that b_N must be some type of logarithmic function. In particular, $b_N = \sigma \log N$ where σ is a constant. We plug this value into Equation (27) to get:

29... $[G(y)]^N = G(y + \sigma \log N)$.

Take the logarithm of both sides of (29) to get :

$$N\{- \log G(y)\} = - \log G(y + \sigma N)$$

where the negative sign emerges from the fact that $G(y) \leq 1$. We take the logarithm once again (sometimes called the *law of iterated logarithms*):

30...$\log N + \log \{- \log G(y)\} = \log \{- \log G(y + \sigma N)\}$.

Let $h(y) = \log \{- \log G(y)\}$. Equation (30) becomes:

31...$\log N + h(y) = h(\sigma (\frac{y}{\sigma} + \log N)$ or $h(y) = h(\sigma (\frac{y}{\sigma} + \log N) - \log N$

Put $\sigma (\frac{y}{\sigma} + \log N) = 0$ (or find h(0) on the right hand side. This means that $\frac{y}{\sigma} = -\log N$. Substituting back to Equation (31), we obtain:

32...$h(y)=h(0)-\frac{y}{\sigma}$, since h(y) decreases as y increases.

Thus, $-\log G(y) = \exp(h(y)) = \exp(h(0)-\frac{y}{\sigma}) = \exp(-(\frac{y-\sigma h(0)}{\sigma})$. Put $\mu = \sigma h(0)$ and we have:

It follows that:

$$33 \ldots -\log G(y) = \exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right).$$

The asymptotic mean and variance of Y can now be obtained from G(y). The mean, variance, skewness and kurtosis are given in the references.

$$34 \ldots G(y) = \exp\left(-\exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right) \text{ or } G(y) = \exp\left(-\exp\left(-\left(\frac{y-\alpha}{\beta}\right)\right)\right).$$

Asymptotic Results for the Independent but Not Identically Distributed Case (INID)

Let $X_1, X_2, \ldots, X_n$ be independent normal random variables with means $E(X_i) = \mu_i$ and variances given by $var(X_i) = \sigma_i^2$. As before, we let $Y = \max\{X_i\}$. The distribution of Y is given by:

$$35 \ldots F_n^*(y) = \Pi_i Fi(y),\ i = 1, 2, \ldots, n.$$

There are two ways in which we can think of the asymptotics or large sample scenario. The first is to increase *n* without bound, in which case we have a sequence of means $\{\mu_i\}$ and variances $\{\sigma_i^2\}$, $i = 1,2,3,\ldots$. The second way is to fix the number n of means and variances and then to increase the number of observations M for each component distribution function i.e take a random sample of size M from each of the *n* component distributions and increase this without bound. In practice, it is the second interpretation that appears to be reasonable and implementable. Hence, our asymptotic analysis will follow the second interpretation.

To this end, let:

$\bar{\mu} = \frac{1}{n}\Sigma\mu_i$ = arithmetic average of the means which is non-stochastic,

$\overline{\sigma^2} = \frac{1}{n}\Sigma\sigma_i^2$ = arithmetic average of the variances also non-stochastic

$\bar{\sigma} = \frac{1}{\sqrt{n}}\sqrt{\Sigma\sigma_i^2}$ = the square root of the arithmetic average of the variances.

We consider:

$$F_n^*\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right) = \Pi_i Fi\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right)],\ i = 1\ 2,\ldots,n.$$

If samples of size M were obtained from each distribution function, then:

$$36 \ldots F_n^*\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right) = \Pi_i\left[Fi\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right)\right]^M,$$ where:

$i = 1\ 2,\ldots,n.,\ M = 1,2,3,\ldots\to\infty$

Assume, as before, that the limiting distribution of (36) exists and is G(.):

$$37 \ldots \Pi_i\left[Fi\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right)\right]^M \to G\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right) \text{ as } M\to\infty.$$

Applying the same *stability postulate* as before, we have that:

$$38 \ldots G\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right)^M = G\left(\frac{y-\bar{\mu}}{\bar{\sigma}} + b_M\right)$$

For which we conclude that $b_M = \theta \log M$. Hence:

$$39 \ldots \log M + \log\left[-\log\left(G\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right)\right)\right] = \log\left(-\log\left(G\left(\frac{y-\bar{\mu}}{\bar{\sigma}} + \theta\log M\right)\right)\right).$$

Equation (39) implies that if $h\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right) = \log\left(-\log\left(G\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right)\right)\right)$, then:

$$40 \ldots h\left(\frac{y-\bar{\mu}}{\bar{\sigma}}\right) = h(0) - \frac{y-\bar{\mu}}{\bar{\sigma}\theta} - \frac{\bar{\sigma}}{\bar{\sigma}\theta}\ h(0))$$

It follows that:

$$41 \ldots G(y) = \exp\left(-\exp\left(-\left(\frac{y-\bar{\mu}-\bar{\sigma}\,h(0)}{\bar{\sigma}\theta}\right)\right)\right)$$

which we recognize as a Type I Gumbel distribution with $\alpha = \bar{\mu} + \bar{\sigma}h(0)$ and $\beta = \bar{\sigma}\theta$. Here $h(0) = \log(-\log(G(0)))$.

References

Abdelkader, Y. (2005). Computing the moments of order statistics from non-identically distributed erlang variables. *Statistical Papers,45,(4), 563-570.*

Arellano-Valle, R.B., Branco, M.D., Genton, M.G. (2006). A unified view on skewed distributions arising from selections. *Canada J. Statist. 34, 581–601.*

Barakat, H. and Abdelkader, Y. (2000). Computing the moments of order statistics from non-Identically distributed Weibull variables. *J. Comp. Appl. Math, 117,(1), 85-90.*

Barakat, H. And Y. Abdelkader. (2004). Computing the moments of order statistics from non-identical random variables. *Statistics Methods & Applications, 13, 15-26).*

David, H.A. (1981). *Order Statistics.* (Wiley, New York).

Galambos, J. (1987). *The asympototic theory of extreme order statistics.* (2nd eds. Krieger, Malabar, Florida).

Genz, A. (1992). Numerical computation of multivariate normal probabilities. *J. Comput. Graph Statist. 1, 141–149.*

Hill, W.G. (1977). Order statistics of correlated variables and implications in genetic selection programmes II. Response to selection. *Biometrics 33, 703–712.*

Padua, R. (1998). Advanced lecture notes on probability and mathematical Statistics. Unpublished monograph, Mindanao Polytechnic State College.