Aymptotics in normal order statistics

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics and Computing Science

Memorandum 1986-05

May 1986

ASYMPTOTICS IN NORMAL ORDER STATISTICS

by

J.J.A.M. Brands

Eindhoven University of Technology
Department of Mathematics and Computing Science
PO Box 513
5600 MB Eindhoven
The Netherlands

ASYMPTOTICS IN NORMAL ORDER STATISTICS

by

J.J.A.M. BRANDS

Department of Mathematics, Eindhoven University of Technology,
The Netherlands

ABSTRACT

In order statistics certain integrals involving the standard normal distribution play an important role. The asymptotic behaviour with respect to a large parameter is studied.

1. INTRODUCTION

The expectation and variance of the maximum in a random sample of size n from the standard normal distribution involve some of the integrals

$$(1) \qquad M_j(n) := \int_{-\infty}^{\infty} x^{j-1}\, \phi(x)(1-\Phi^{n-1}(x))\,dx \qquad (n \in \mathbb{N}, j \in \mathbb{N}) \;;$$

where

$$(2) \qquad \Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-s^2/2}\, ds \;.$$

Integration by parts gives

$$(3) \qquad \mu_j(n) = j\,M_j(n) + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^j\, e^{-x^2/2}\, dx \;,$$

where

$$(4) \qquad \mu_j(n) := \frac{n}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^j\, e^{-x^2/2}\, \Phi^{n-1}(x)\,dx \qquad (n \in \mathbb{N}, j \in \mathbb{N}) \;.$$

The problem posed by two colleagues *) of the author is to determine the asymptotic behaviour for $n \to \infty$ of $\mu_j(n)$ for $j = 1,2,3,4$. Moreover, they are interested especially in the asymptotic behaviour of

$$(\mu_3 - \mu_1\mu_2)(\mu_2 - \mu_1^2)^{-\frac{1}{2}} \quad \text{and} \quad (\mu_4 - \mu_2^2)^{\frac{1}{2}} .$$

We remark that the differences

$$M_j(n+1) - M_j(n) = \int_{-\infty}^{\infty} x^{j-1}(1-\Phi(x))\Phi^n(x)\,dx$$

are just the integrals occurring in the coefficients of the asymptotic formulas in a previous paper [1] (where f is defined by $f(x) = \Phi(x\sqrt{2})$).

2. RESULTS

Let the asymptotic series of $(1-\Phi(x))(\Phi'(x))^{-1}$ by denoted by A, i.e.

$$A := \sum_{\ell=0}^{\infty} (-1)^{\ell}(2\ell-1)!!\, x^{-2\ell-1} \qquad (x \to \infty) \;, \quad ((-1)!! = 1) \;.$$

Formal differentiations of A are denoted by A', A'', etc. Let $x_1 = x_1(n)$ be defined by $\Phi(x_1) = 1 - \frac{1}{n}$. Then $\mu_j(n)$ has an asymptotic expansion in powers of x_1^{-1}, i.e.

$$\mu_j(n) \approx \sum_{k=-j}^{\infty} C(j,k)x_1^{-k} \qquad (n \to \infty) \;,$$

which can be computed as follows: x is considered to be a function of a variable z and $\frac{dx}{dz} \approx \frac{1}{2}A$. Higher derivatives can be computed by means of the chain rule. For instance, $\frac{d^2x}{dz^2} \approx \frac{1}{4}AA'$. Then

$$\mu_j(n) \approx \sum_{\ell=0}^{\infty} \left(\frac{d^{\ell}\, x^j}{dz^{\ell}}\right)_1 \frac{(-2)^{\ell}\, \Gamma^{(\ell)}(1)}{\ell!} \qquad (n \to \infty) \;,$$

where the subscript 1 means that the value at $x = x_1$ has to be taken. x_1 has the following asymptotic series:

---

*) F.W. Steutel and D.A. Overdijk, Department of Mathematics, Eindhoven University of Technology, The Netherlands.

$$x_1 \approx z^{\frac{1}{2}}\left(1 + \sum_{k=1}^{\infty} q_k(\log z) z^{-k}\right) \qquad (n \to \infty)$$

where $z = 2 \log \frac{n}{\sqrt{2\pi}}$ and the q_k's are polynomials of degree k. A few q_k's are

$$q_1(t) = -\frac{1}{2}t, \quad q_2(t) = -\frac{1}{8}t^2 + \tfrac{1}{2}t - 1,$$

$$q_3(t) = -\frac{1}{16}t^3 + \frac{1}{2}t^2 - \frac{3}{2}t + \frac{7}{2}.$$

The coefficients $C(j,k)$ have the property that $C(j,-j+s) = 0$ if s is odd and $C(j,-j) = 1$. A few more coefficients $C(j,k)$ are:

$$C(1,1) = -\Gamma'(1), \quad C(1,3) = \Gamma'(1) - \frac{1}{2}\Gamma''(1),$$

$$C(1,5) = 3\Gamma'(1) + 2\Gamma''(1) - \frac{1}{2}\Gamma'''(1);$$

$$C(2,0) = -2\Gamma'(1), \quad C(2,2) = 2\Gamma'(1), \quad C(2,4) = -6\Gamma'(1) + 2\Gamma''(1),$$

$$C(2,6) = 30\Gamma'(1) - 14\Gamma''(1) + \frac{8}{3}\Gamma'''(1);$$

$$C(3,-1) = -3\Gamma'(1), \quad C(3,1) = 3\Gamma'(1) + \frac{3}{2}\Gamma''(1),$$

$$C(3,3) = -9\Gamma'(1) + \frac{1}{2}\Gamma'''(1),$$

$$C(3,5) = 45\Gamma'(1) - \frac{21}{2}\Gamma''(1) - \frac{1}{2}\Gamma'''(1) + \frac{3}{8}\Gamma^{(4)}(1);$$

$$C(4,-2) = -4\Gamma'(1), \quad C(4,0) = 4\Gamma'(1) + 4\Gamma''(1),$$

$$C(4,2) = -12\Gamma'(1) - 4\Gamma''(1), \quad C(4,4) = 60\Gamma'(1) - \frac{8}{3}\Gamma'''(1).$$

A routine computation shows that

$$\frac{\mu_3 - \mu_1\mu_2}{(\mu_2 - \mu_1^2)^{\frac{1}{2}}} \approx d_0 + d_2 x_1^{-2} + d_4 x_1^{-4} + \ldots \qquad (n \to \infty)$$

$$(\mu_4 - \mu_2^2)^{\frac{1}{2}} \approx e_0 + e_2 x_1^{-2} + e_4 x_1^{-4} + \ldots \qquad (n \to \infty)$$

where

$$d_0 = e_0 = 2(\Gamma''(1) - (\Gamma'(1))^2)^{\frac{1}{2}}$$

and

$$d_2 = e_2 = - d_0 .$$

3. PROOF OF THE RESULTS

We transform the integral in (4) by putting

(5) $$\Phi(x) = 1 - \frac{s}{n} .$$

Then

(6) $$\frac{dx}{ds} = - \frac{\sqrt{2\pi}}{n} e^{x^2/2} ,$$

whence

(7) $$\mu_j(n) = \int_0^n x^j (1-\frac{s}{n})^{n-1} \, ds .$$

We observe that $x = x(s)$ is monotonically decreasing on $[0,\infty)$, that $x(s) \to \infty$ $(s \downarrow 0)$, $x(\frac{1}{2}n) = 0$ and $x(s) \to -\infty$ $(s \uparrow n)$.

Now we shall prove that

(8) $$\mu_j(n) = \left(1 + \mathcal{O}(\frac{\log^2 n}{n})\right) \int_0^{\log n} x^j e^{-s} \, ds + \mathcal{O}(n^{-1}(\log n)^{j/2}) \qquad (n \to \infty) .$$

Since

$$\left| \int_{\pi/2}^{n} x^j (1-\frac{s}{n})^{n-1} \, ds \right| = \frac{n}{\sqrt{2\pi}} \int_{-\infty}^{0} |x|^j \, e^{-x^2/2} \, \Phi^{n-1}(x) dx \leq$$

$$\leq \frac{n2^{-n}}{\sqrt{2\pi}} \int_0^{\infty} x^j \, e^{-x^2/2} \, dx$$

we can write

(9) $$\mu_j(n) = \int_0^{\frac{1}{2}n} x^j (1-\frac{s}{n})^{n-1} \, ds + \mathcal{O}(n2^{-n}) \qquad (n \to \infty) .$$

Using

$$(10) \qquad \Phi(x) \approx 1 - \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \sum_{\ell=0}^{\infty} (-1)^{\ell}(2\ \ -1)!!\ x^{-2\ell-1} \qquad (x \to \infty)$$

we derive easily that at $s = \log n$

$$(11) \qquad x \sim \sqrt{2 \log n} \qquad (n \to \infty) .$$

Hence

$$(12) \qquad \int_{\log n}^{n/2} x^j (1-\tfrac{s}{n})^{n-1}\, ds = \mathcal{O}\Big((\log n)^{j/2} \int_{\log n}^{n/2} (1-\tfrac{s}{n})^{n-1}\, ds\Big) =$$

$$= \mathcal{O}(n^{-1}(\log n)^{j/2}) \qquad (n \to \infty) .$$

From (9) and (12) it follows that

$$(13) \qquad \mu_j(n) = \int_0^{\log n} x^j (1-\tfrac{s}{n})^{n-1}\, ds + \mathcal{O}(n^{-1}(\log n)^{j/2}) \qquad (n \to \infty) .$$

Furthermore

$$(14) \qquad \int_0^{\log n} x^j (1-\tfrac{s}{n})^{n-1}\, ds = \Big(1 + \mathcal{O}(\tfrac{\log^2 n}{n})\Big) \int_0^{\log n} x^j e^{-s}\, ds \qquad (n \to \infty)$$

since

$$(15) \qquad (1-\tfrac{s}{n})^{n-1} = e^{-s}(1 + \mathcal{O}(\tfrac{\log^2 n}{n})) \qquad (0 \leq s \leq \log n,\ n \to \infty) .$$

Clearly (13) and (14) imply (8).

On the interval $0 \leq s \leq \log n$, corresponding with large values of x, we can use (10) in order to solve x from (5) as a function of s.
Introduction of

$$(16) \qquad z = 2 \log \frac{n}{\sqrt{2\pi}} - 2 \log s$$

transforms (5) into

$$(17) \qquad \Phi(x) = 1 - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z} .$$

Clearly $z \to \infty$ if $n \to \infty$ and $0 < s < \log n$.
Using (10) and taking logarithms we get

$$(18) \qquad z \approx x^2 + \log(x^2) + \frac{2}{x^2} - \frac{5}{x^4} + \frac{74/3}{x^6} \ldots \qquad (x \to \infty)$$

By asymptotic iteration we find

$$(19) \qquad x^2 \approx z - \log z + \sum_{k=1}^{\infty} z^{-k} p_k(\log z) \qquad (z \to \infty)$$

where the p_k's are polynomials of degree k. A few polynomials p_k are

$$(20) \qquad p_1(t) = t - 2$$

$$p_2(t) = \frac{1}{2} t^2 - 3t + 7$$

$$p_3(t) = \frac{1}{3} t^3 - \frac{3}{2} t^2 + 17t - \frac{107}{3} .$$

The asymptotic expansions for x^j have the form

$$(21) \qquad x^j \approx z^{j/2}\left(1 + \sum_{k=1}^{\infty} p_{jk}(\log z) z^{-k}\right) \qquad (z \to \infty)$$

where the p_{jk} are polynomials of degree k.
The individual terms in the asymptotic expansions (21) have the following property: Let $f(z)$ be such a term occurring in the right side of (21). Then $f(z)$ is of the form

$$f(z) = (\log^m z) z^{-k+\frac{1}{2}j} \qquad \text{where } m \leq k .$$

Let $z_1 := 2 \log \frac{n}{\sqrt{2\pi}}$. Clearly $z_1 \geq 2$ if $n \geq 7$. Let $n \geq 7$. Then the powerseries about $z = z_1$

$$(22) \qquad f(z_1 + \varepsilon) = \sum_{\ell=0}^{\infty} \frac{f^{(k)}(z_1)}{\ell!} \varepsilon^{\ell}$$

is convergent for $|\varepsilon| < z_1$. Now it is easily seen that this powerseries has the property that for all $N \in \mathbb{N}$, $N \geq \frac{1}{2}j-k$

$$(23) \qquad f(z_1 + \varepsilon) = \sum_{\ell=0}^{N} \frac{f^{(\ell)}(z_1)}{\ell!} \varepsilon^{\ell} + \mathcal{O}((\log^m z_1) z_1^{-k+\frac{1}{2}j-N-1} \varepsilon^{N+1})$$

$$(-\tfrac{1}{2}z_1 < \varepsilon < \infty,\ z_1 \geq 2) .$$

Then it follows that upon substitution $\varepsilon = -2\log s$ in (22) we get an asymptotic expansion for $0 < s < \log n$, $n \to \infty$, i.e. for all $N \geq \frac{1}{2}j-k$

$$(24) \qquad f(z_1 - 2\log s) = \sum_{\ell=0}^{N} \frac{f^{(\ell)}(z_1)}{\ell!}(-2\log s)^{\ell} +$$

$$+ \mathcal{O}((\log^m z_1) z_1^{-k+\frac{1}{2}j-N-1}(\log s)^{N+1}) \qquad (0 < s < \log n,\ n \to \infty)$$

since $2\log\log n < \frac{1}{2}z_1$ for n sufficiently large. The hidden constant in the O-term is independent of n.

Further, for every $\ell \in \mathbb{N}$,

$$(25) \qquad \int_{\log n}^{\infty} \log^{\ell} s\, e^{-s}\, ds = \mathcal{O}(n^{-1}(\log\log n)^{\ell}) \qquad (n \to \infty)\ .$$

Therefore we can proceed as follows: In the asymptotic expansion (21) of x^j we substitute $z = z_1 - 2\log s$ and we expand formally into a power-series about z_1. After multiplication with e^{-s} and integration over $(0,\infty)$ we get an asymptotic expansion for $\mu_j(n)$.

Denoting the asymptotic expansion (21) of x^j by X_j and writing $X_j^{(k)}$ for its formal derivatives we have proved that

$$(26) \qquad \mu_j(n) \approx \sum_{k=0}^{n} X_j^{(k)}(z_1)(-2)^k\, \Gamma^{(k)}(1)(k1)^{-1} \qquad (n \to \infty)\ ,$$

where we have used that

$$(27) \qquad \int_0^{\infty} e^{-s} \log^k s\, ds = \Gamma^{(\ell)}(1)\ .$$

If we carry out the above program then, for instance, we find

$$(28) \qquad \mu_1 = z^{1/2} - \frac{1}{2} z^{-1/2}\log z + \gamma z^{-1/2} - \frac{1}{8} z^{-3/2}\log^2 z +$$

$$+ \frac{1}{2}(1+\gamma) z^{-3/2}\log z - (1-\gamma-\frac{1}{2}\gamma^2 - \frac{1}{12}\pi^2) z^{-3/2} +$$

$$+ \mathcal{O}(z^{-5/2}\log^3 z) \qquad (n \to \infty)$$

where $z = z_1$. We have used that $\Gamma'(1) = \gamma$ (Euler's constant) and $\Gamma''(1) = \gamma^2 + \frac{1}{6}\pi^2$.

Of course we can also find asymptotic results for μ_2, μ_3 and μ_4. We will not do so since there is a more convenient way to obtain asymptotic expansions for $\mu_j(n)$. We shall show that $\mu_j(n)$ has an asymptotic power-series expansion in powers of x_1^{-1}, where x_1 is the value of x at $z = z_1 := 2 \log \frac{n}{\sqrt{2\pi}}$, i.e. there are sequences $(C(j,k))_{k=-j}^{\infty}$ of real numbers such that

$$(29) \qquad \mu_j(n) \approx \sum_{k=-j}^{\infty} C(j,k)x_1^{-k} \qquad (n \to \infty) .$$

Considering x as a function of z defined by (17) we can write the integral in (8) as

$$(30) \qquad \int_0^{\log n} x^j(z_1 - 2 \log s)e^{-s}\, ds .$$

We shall prove that we can find the asymtpotic expansion of (30) by termwise integration of the formal powerseries expansion of $x^j(z_1 - 2 \log s)$ about z_1. We shall give the details of the proof for the case $j = 1$. The other cases $j > 1$ can be treated analogously. So for the moment being we suppose $j = 1$. Obviously we are done with the problem if we have proved that

$$(31) \qquad \forall_{k \in \mathbb{N}} \left(\frac{d^k x}{dz^k}\right)_1 \text{ has an asymptotic powerseries in } x_1^{-1} .$$

$$(32) \qquad \forall_{k \in \mathbb{N}} \left(\frac{d^{k+1} x}{dz^{k+1}}\right)_1 = \mathcal{O}\left(\left(\frac{d^k x}{dz^k}\right)_1\right) \qquad (n \to \infty)$$

$$(33) \qquad \forall_{N \in \mathbb{N}} \exists_{A>0} \; x(z_1 - 2 \log s) = x_1 + \sum_{k=1}^{N} \left(\frac{d^k x}{dz^k}\right)_1 \frac{(-2 \log s)^k}{k!} +$$

$$+ \mathcal{O}\left(\left(\frac{d^{N+1} x}{dz^{N+1}}\right)_1 (\log s)^{N+1}\right) \qquad (0 < s < \log n,\ n > A)$$

where the subscript 1 in $(\ )_1$ means the value at z_1.

From (17) it follows that

$$(34) \qquad \frac{dx}{dz} = \tfrac{1}{2}a(x) \ ,$$

where

$$(35) \qquad a(x) = \frac{1 - \Phi(x)}{\Phi'(x)} \ .$$

From (10) we see that

$$(36) \qquad a(x) \approx \sum_{\ell=0}^{\infty} (-1)^{\ell}(2\ell-1)!! \ x^{-2\ell-1} \qquad (x \to \infty) \ .$$

From (35) we derive that

$$(37) \qquad \frac{da}{dx} = xa - 1 \ .$$

Clearly (36) and (37) imply that all derivatives $d^k a/dx^k$ have asymptotic powerseries in x^{-1} which can be obtained by formal differentiation of the asymptotic series in (36). By means of the chain rule we can compute from (34) all derivatives dx^k/dz^k; clearly, dx^k/dz^k is a sum of products involving $a(x)$ and its derivatives $d^{\ell}a/dx^{\ell}$, $\ell = 1,2,\dots,k-1$. It follows that $d^k x/dz^k$ has an asymptotic expansion in powers of x^{-1} for $x \to \infty$. Using that

$$\frac{d^{\ell}a}{dx^{\ell}} \sim \frac{(-1)^{\ell} \ \ell!}{x^{\ell+1}} \qquad (x \to \infty)$$

we easily derive that for $k \geqq 2$

$$(38) \qquad \frac{d^k x}{dz^k} \sim (-1)^{k+1}(2k-3)!! \ 2^{-k} \ x^{-2k+1} \qquad (x \to \infty) \ .$$

Thus we have proved (31) and (32).

Let $N \in \mathbb{N}$. Let $h \in \mathbb{R}$. Then there is a number $\theta \in (0,1)$ such that

$$(39) \qquad x(z_1 + h) = x_1 + \frac{h}{1!}\left(\frac{dx}{dz}\right)_1 + \dots + \frac{h^N}{N!}\left(\frac{d^N x}{dz^N}\right)_1 + R_N$$

where

$$(40) \qquad R_N = \frac{h^{N+1}}{(N+1)!} \frac{d^{N+1}x}{dz^{N+1}} (z_1 + \theta h) \ .$$

Restricting ourselves to $h > -\frac{1}{2}z_1$, we only have to prove that there is a number $A > 0$ and a number $K > 0$ such that for all $z > A$

$$(41) \qquad \left|\frac{d^{N+1}x}{dz^{N+1}}(\eta)\right| < K\left|\frac{d^{N+1}x}{dz^{N+1}}(z)\right| \qquad (\eta > \tfrac{1}{2}z) .$$

On account of (38) and the fact that $x(z) \sim z^{\frac{1}{2}}$ $(z \to \infty)$, (41) is obviously true. Hence (33) holds, since $z_1 - 2\log s > \frac{1}{2}z_1$ for $0 < s < \log n$ and n sufficiently large.

Now using (33) in (30) for $j = 1$ we get an asymptotic series in powers of x_1^{-1} for $\mu_1(n)$.
Analogously we can find asymptotic series for $\mu_j(n)$, $j = 2,3,\ldots$. Only small adaptations are necessary; for instance in (33) we have to change $\forall_{N\in\mathbb{N}}$ into $\forall_{N\in\mathbb{N},N\geq\frac{1}{2}j}$.

REMARK. Especially for $(\mu_3 - \mu_1\mu_2)(\mu_2 - \mu_1^2)^{-\frac{1}{2}}$ and $(\mu_4 - \mu^2)^{\frac{1}{2}}$ the computations are most easily done if one postpones the replacement of $d^\ell x^j/dz^\ell$ by its asymptotic series as long as possible. For instance, in this way we get

$$\mu_4 - \mu_2^2 = (8B - 4A^2)x^2(x')^2 + (24C - 8AB)(x^2x'x'' + x(x')^3) + \mathcal{O}(x^4)$$

where

$$A = -2\Gamma'(1) \ , \quad B = 2\Gamma''(1) \quad \text{and} \quad C = -\frac{4}{3}\Gamma'''(1) \ ,$$

and x', x'' denote first and second derivatives with respect to z. Now using the asymptotic series

$$x' = \frac{1}{2x} - \frac{1}{2x^3} + \ldots \ , \qquad x'' = -\frac{1}{4x^3} + \frac{1}{x^5} + \ldots$$

we see that

$$x^2x'x'' + x(x')^3 = \mathcal{O}(x^{-4}) \quad \text{and} \quad x^2(x')^2 = \frac{1}{4} - \frac{1}{2x^2} + \ldots \ .$$

We get

$$\mu_4 - \mu_2^2 = d_0^2 - 2d_0^2x^{-2} + \mathcal{O}(x^{-4}) \ ,$$

from which we can easily derive the result in section 2.

REFERENCE

1 BRANDS, J.J.A.M., Asymptotics in Poisson order statistics, Memorandum 86-03, Department of Mathematics, Eindhoven University of Technology, The Netherlands.