TOPIC The normalizing transformation and friends. This section studies some asymptotic expansions that are closely related to Edgeworth's expansion for the cdf.

Let X be a real random variable with a continuous strictly increasing distribution function F. Let Z be a standard normal random variable, with distribution function Φ. For real numbers x, consider

$$N(x) := \Phi^{-1}\big(F(x)\big). \tag{1}$$

Note that

$$P[\, Z \le N(x)\,] = \Phi\big(\Phi^{-1}(F(x))\big) = F(x) = P[\, X \le x\,],$$

so $N(x)$ is the same quantile of Z that x is of X. For example, when $X \sim \chi^2_{12}$, we have

$x =$	17	32	53
$N(x) =$	1	3	5

In general, $N(x)$ is called ***the normal deviate equivalent to x***, or just the ***equivalent normal deviate***. Since (3) below implies

$$N(X) = \Phi^{-1}\big(F(X)\big) \sim Z,$$

N is called the ***normalizing transformation***. Note that it is continuous and strictly increasing.

The inverse C to N is also interesting:

$$C(z) := N^{-1}(z) = F^{-1}\big(\Phi(z)\big). \tag{2}$$

Evidently

$$P[\, X \le C(z)\,] = F\big(F^{-1}(\Phi(z))\big) = \Phi(z) = P[\, Z \le z\,], \tag{3}$$

so $C(z)$ is the same quantile for X that z is for Z. For example the 0.975 quantile of X is $C(1.96)$. Since $C(Z) \sim X$, C is called the ***inverse normalizing***, or ***Cornish-Fisher***, ***transformation.***

Now suppose that X has mean 0, standard deviation 1, and is almost normally distributed. Then the normalizing transformation and

its inverse will both be nearly equal to the identity transformation I. If we know how X departs from normality, we should be able to say how N and C differ from I. To be specific, suppose that X is indexed by an integer $n \in \mathbb{N}$ and that as n tends to infinity, the distribution function F_n of X_n has a second-order Edgeworth approximation of the form

$$F_n(x) = \Phi(x) - \phi(x)\Omega^{(2)}_{n,E}(x) + \rho^{(2)}_{n,E}(x) \tag{4}$$

where $\phi = \Phi'$,

$$\Omega^{(2)}_{n,E}(x) = \Big[\frac{\kappa_3 H_2(x)}{6}\Big]\frac{1}{\sqrt{n}} + \Big[\frac{\kappa_4 H_3(x)}{24} + \frac{\kappa_3^2 H_5(x)}{72}\Big]\frac{1}{n}, \tag{5}$$

and the remainder term $\rho^{(2)}_{n,E}(x)$ is $O(1/n^{3/2})$ locally uniformly in x, i.e., for each $x \in \mathbb{R}$, there exists an $\epsilon > 0$ such that

$$\sup\{\, |\rho^{(2)}_{n,E}(\xi)| : |\xi - x| \le \epsilon \,\} = O\Big(\frac{1}{n^{3/2}}\Big). \tag{6}$$

We are going to establish analogous approximations for

$$N_n := \Phi^{-1}(F_n) \quad \text{and} \quad C_n := F_n^{-1}(\Phi).$$

For simplicity of expression, we write "O^*" to mean "***O, locally uniformly***". The approximation for N_n follows from this result:

Theorem 1. *Let f_1, f_2, ... be an infinite sequence of functions having an asymptotic expansion of the form*

$$f_n(x) = c_0(x) + c_1(x)b_n + c_2(x)b_n^2 + O^*(b_n^3) \tag{7}$$

where c_0, c_1, and c_2 are bounded on bounded intervals and $b_n = o(1)$. Let g be a three-times continuously differentiable function defined on the range of the f_n's. Then the composite functions $g(f_n)$ have the asymptotic expansion

$$\begin{aligned} g\big(f_n(x)\big) &= g\big(c_0(x)\big) + \big[g'\big(c_0(x)\big)c_1(x)\big]b_n \\ &\quad + \big[g'\big(c_0(x)\big)c_2(x) + g''\big(c_0(x)\big)c_1^2(x)/2\big]b_n^2 + O^*(b_n^3). \end{aligned} \tag{8}$$

"O^*" means "O, locally uniformly".

(7): $f_n(x) = c_0(x) + c_1(x)b_n + c_2(x)b_n^2 + O^*(b_n^3)$.

Proof Put $h_n(x) = g\big(f_n(x)\big)$ and

$$\hat{h}_n(x) = g\big(c_0(x)\big) + \big[g'\big(c_0(x)\big)c_1(x)\big]b_n + \big[g'\big(c_0(x)\big)c_2(x) + g''\big(c_0(x)\big)c_1^2(x)/2\big]b_n^2.$$

(8) asserts that $h_n(x) = \hat{h}_n(x) + O^*(b_n^3)$, or, equivalently — see Exercise 1 — that

$$h_n(x_n) = \hat{h}_n(x_n) + O(b_n^3) \tag{9}$$

for every convergent sequence $(x_n)_{n=1}^\infty$. To verify (9), suppose that x_n converges, say to x. We are first going to show that

$$h_n(x_n) = g\big(\varphi_n(x_n)\big) + O(b_n^3) \tag{10}$$

where

$$\varphi_n(\xi) = c_0(\xi) + c_1(\xi)b_n + c_2(\xi)b_n^2$$

for $\xi \in \mathbb{R}$. By the mean value theorem,

$$h_n(x_n) = g\big(f_n(x_n)\big) = g\big(\varphi_n(x_n)\big) + g'\big(y_n\big)\rho_n$$

where $\rho_n = f_n(x_n) - \varphi_n(x_n)$ and y_n lies between $\varphi_n(x_n)$ and $f_n(x_n)$. We have $\rho_n = O(b_n^3)$ by (7). Since c_0, c_1, and c_2 are bounded on bounded intervals, the sequence $(\varphi_n(x_n))$ is bounded, and therefore so is the sequence (y_n), since y_n lies between $\varphi_n(x_n)$ and $\varphi_n(x_n)+\rho_n$. Since g' is bounded on bounded intervals, it follows that (10) holds. Now write

$$\varphi_n(x_n) = c_0(x_n) + \eta_n \quad \text{with} \quad \eta_n = c_1(x_n)b_n + c_2(x_n)b_n^2.$$

Note that $\eta_n = O(b_n)$. Arguing as above, but using a second order Taylor expansion with remainder in place of the MVT, we get

$$g\big(\varphi_n(x_n)\big) = g\big(c_0(x_n) + \eta_n\big) = g\big(c_0(x_n)\big) + g'\big(c_0(x_n)\big)\eta_n + g''\big(c_0(x_n)\big)\eta_n^2/2 + O(b_n^3) = \hat{h}_n(x_n) + O(b_n^3), \tag{11}$$

the last step using $\eta_n^2 = c_1^2(x_n)b_n^2 + O(b_n^3)$. (10) and (11) give (9). ∎

(4): $F_n(x) = \Phi(x) - \phi(x)\Omega_{n,E}^{(2)}(x) + \rho_{n,E}^{(2)}(x)$.

(7): $f_n(x) = c_0(x) + c_1(x)b_n + c_2(x)b_n^2 + O^*(b_n^3)$.

Theorem 2. *Suppose* (4) *holds with* $\Omega_{n,E}^{(2)}$ *defined by* (5) *and the remainder* $\rho_{n,E}^{(2)}(x) = O(1/n^{3/2})$ *locally uniformly in* x. *Then*

$$N_n(x) = x - \Omega_{n,N}^{(2)}(x) + O^*\Big(\frac{1}{n^{3/2}}\Big) \tag{12}$$

where

$$\Omega_{n,N}^{(2)}(x) = \Big[\frac{\kappa_3 H_2(x)}{6}\Big]\frac{1}{\sqrt{n}} + \Big[\frac{\kappa_4 H_3(x)}{24} + \frac{\kappa_3^2\big(-4H_3(x) - 5H_1(x)\big)}{36}\Big]\frac{1}{n}. \tag{13}$$

$N_n^{(2)}(x) := x - \Omega_{n,N}^{(2)}(x)$ is called the second-order approximation to the normalizing transformation. Loosely stated Theorem 2 says that the normal deviate equivalent to x is nearly $N_n^{(2)}(x)$ and hence that $N_n^{(2)}(X_n)$ is a polynomial transformation of X_n that is almost standard normal. Note that for large n, $N_n^{(2)}$ is strictly increasing over an interval containing most of the mass of the distribution of X_n; however, for any n, it behaves badly for extreme values of X_n.

Proof of Theorem 2. Theorem 1 with

$$\begin{aligned}
&b_n = 1/\sqrt{n},\\
&f_n(x) = F_n(x),\\
&c_0(x) = \Phi(x),\\
&c_1(x) = \phi(x)C_1(x) \text{ for } C_1(x) = -\kappa_3 H_2(x)/6,\\
&c_2(x) = \phi(x)C_2(x) \text{ for } C_2(x) = -\big(\kappa_4 H_3(x)/24 + \kappa_3^2 H_5(x)/72\big),\\
&g(u) = R(u) := \Phi^{-1}(u)
\end{aligned}$$

$c_0(x) = \Phi(x)$. $\quad c_1(x) = \phi(x)C_1(x)$. $\quad c_2(x) = \phi(x)C_2(x)$.
$C_1(x) = -\kappa_3 H_2(x)/6$. $\quad C_2(x) = -\big(\kappa_4 H_3(x)/24 + \kappa_3^2 H_5(x)/72\big)$.
$b_n = 1/\sqrt{n}$. $\quad g = R := \Phi^{-1}$.

gives

$$\begin{aligned}
N_n(x) &= \Phi^{-1}\big(F_n(x)\big) = g\big(f_n(x)\big) \\
&= g\big(c_0(x)\big) + \big[g'\big(c_0(x)\big)c_1(x)\big]b_n \\
&\quad + \big[g'\big(c_0(x)\big)c_2(x) + g''\big(c_0(x)\big)c_1^2(x)/2\big]b_n^2 + O^*(b_n^3) \\
&= R\big(\Phi(x)\big) + \big[R'\big(\Phi(x)\big)\phi(x)C_1(x)\big]b_n \\
&\quad + \big[R'\big(\Phi(x)\big)\phi(x)C_2(x) + R''\big(\Phi(x)\big)\phi^2(x)C_1^2(x)/2\big]b_n^2 \\
&\quad + O^*(b_n^3). \qquad (14)
\end{aligned}$$

We need to evaluate $R^{(j)}\big(\Phi(x)\big)\phi^j(x)$ for $j = 0$, 1, and 2. Since $R = \Phi^{-1}$ we have

$$x = R\big(\Phi(x)\big).$$

Differentiating this with respect to x gives

$$1 = R'\big(\Phi(x)\big)\phi(x).$$

Differentiating again gives

$$\begin{aligned}
0 &= R''\big(\Phi(x)\big)\phi^2(x) + R'\big(\Phi(x)\big)\phi'(x) \\
&= R''\big(\Phi(x)\big)\phi^2(x) - R'\big(\Phi(x)\big)\phi(x)x = R''\big(\Phi(x)\big)\phi^2(x) - x.
\end{aligned}$$

Substituting these identities into (14) gives

$$\begin{aligned}
N_n(x) &= x + C_1(x)b_n + [C_2(x) + xC_1^2(x)/2]b_n^2 + O^*(b_n^3) \\
&= x - \Omega_{n,N}^{(2)}(x) + O^*(b_n^3),
\end{aligned}$$

thereby establishing (12). ∎

The asymptotic expansion (12) for N_n has the form

$$f_n(x) = x + c_1(x)b_n + c_2(x)b_n^2 + O^*(b_n^3) \qquad (15)$$

where f_n is continuous and strictly increasing, c_1 and c_2 are smooth functions of x, and $b_n = o(1)$. We are going to show that in this situation there are function d_1 and d_2 of y such that the inverse g_n to f_n satisfies

$$g_n(y) = y + d_1(y)b_n + d_2(y)b_n^2 + O^*(b_n^3). \qquad (16)$$

Since $C_n = N_n^{-1}$, this result and Theorem 2 will immediately yield a second-order approximation to C_n.

Here is the idea behind (16). Loosely stated, (15) reads

$$f_n(x) \approx \varphi_n(x) := x + c_1(x)b_n + c_2(x)b_n^2.$$

Substitute

$$y + d_1(y)b_n + d_2(y)b_n^2 := y + \Delta_n(y)$$

for x, expand c_1 and c_2 in Taylor series, and drop negligible terms to get

$$\begin{aligned}
f_n\big(y + \Delta_n(y)\big) &\approx \varphi_n\big(y + \Delta_n(y)\big) \\
&= \big(y + \Delta_n(y)\big) + c_1\big(y + \Delta_n(y)\big)b_n + c_2\big(y + \Delta_n(y)\big)b_n^2 \\
&\approx \big(y + \Delta_n(y)\big) + [c_1(y) + c_1'(y)\Delta_n(y)]b_n + c_2(y)b_n^2 \\
&\approx y + [d_1(y) + c_1(y)]b_n + [d_2(y) + c_1'(y)d_1(y) + c_2(y)]b_n^2.
\end{aligned}$$

Thus

$$f_n\big(y + \Delta_n(y)\big) \approx y \quad \text{and so} \quad g_n(y) \approx y + \Delta_n(y)$$

provided we take

$$\begin{aligned}
d_1(y) &= -c_1(y) \\
d_2(y) &= -[c_1'(y)d_1(y) + c_2(y)] = c_1(y)c_1'(y) - c_2(y).
\end{aligned}$$

Theorem 3. *Suppose $f_1, f_2, \ldots$ is an infinite sequence of strictly increasing continuous functions having an asymptotic expansion of the form*

$$f_n(x) = x + c_1(x)b_n + c_2(x)b_n^2 + O^*(b_n^3) \tag{17}$$

where c_1 is twice continuously differentiable, c_2 is continuously differentiable, and $b_n = o(1)$. Then the inverse functions $g_n = f_n^{-1}$ have the asymptotic expansion

$$g_n(y) = y + d_1(y)b_n + d_2(y)b_n^2 + O^*(b_n^3) \tag{18}$$

where

$$d_1(y) = -c_1(y) \quad \text{and} \quad d_2(y) = c_1(y)c_1'(y) - c_2(y). \tag{19}$$

Warning: The version of this theorem where the O^*'s in (17) and (18) are replaced by O's is false; see Exercise 4.

Proof of Theorem 3. For arbitrary x and y set

$$\varphi_n(x) = x + c_1(x)b_n + c_2(x)b_n^2$$
$$\gamma_n(y) = y + \Delta_n(y) \quad \text{with} \quad \Delta_n(y) = d_1(y)b_n + d_2(y)b_n^2.$$

To establish (18) it suffices to show that

$$g_n(y_n) - \gamma_n(y_n) = O(b_n^3) \tag{20}$$

for any convergent sequence (y_n). Suppose then that $y_n \to y_0$. To get (20), note that since φ_n is differentiable, the mean value theorem gives

$$\varphi_n\big(g_n(y_n)\big) - \varphi_n\big(\gamma_n(y_n)\big) = \big(g_n(y_n) - \gamma_n(y_n)\big)\varphi_n'(x_n)$$

for some point x_n between $g_n(y_n)$ and $\gamma_n(y_n)$. I will show that

$$\varphi_n\big(g_n(y_n)\big) = y_n + O(b_n^3) \tag{21}$$
$$\varphi_n\big(\gamma_n(y_n)\big) = y_n + O(b_n^3) \tag{22}$$

(so that $\varphi_n\big(g_n(y_n)\big) - \varphi_n\big(\gamma_n(y_n)\big) = O(b_n^3)$) and that

$$\varphi_n'(x_n) = 1 + O(b_n); \tag{23}$$

this will give (20).

(17): $f_n(x) = \varphi_n(x) + O^*(b_n^3)$. $\quad$ $\varphi_n(x) = x + c_1(x)b_n + c_2(x)b_n^2$.
$\gamma_n(y) = y + \Delta_n(y)$ $\quad$ $\Delta_n(y) = d_1(y)b_n + d_2(y)b_n^2$.
(21): $\varphi_n\big(g_n(y_n)\big) = y_n + O(b_n^3)$. $\quad$ (22): $\varphi_n\big(\gamma_n(y_n)\big) = y_n + O(b_n^3)$
(23): $\varphi_n'(x_n) = 1 + O(b_n)$.

• (22) *holds*. Since d_1 and d_2 are continuous, we have $\eta_n := \Delta_n(y_n) = O(b_n)$. This and the smoothness assumptions on c_1 and c_2 imply that

$$\begin{aligned}\varphi_n\big(\gamma_n(y_n)\big) &= \big(y_n + \eta_n\big) + c_1\big(y_n + \eta_n\big)b_n + c_2\big(y_n + \eta_n\big)b_n^2 \\ &= \big(y_n + \eta_n\big) + [c_1(y_n) + c_1'(y_n)\eta_n]b_n + c_2(y_n)b_n^2 + O(b_n^3) \\ &= y_n + [d_1(y_n) + c_1(y_n)]b_n \\ &\quad + [d_2(y_n) + c_1'(y_n)d_1(y_n) + c_2(y_n)]b_n^2 + O(b_n^3),\end{aligned}$$

so (22) holds by the choice of d_1 and d_2.

• (21) *holds*. Recall that $y_n \to y_0$. Since

$$\varphi_n\big(g_n(y_n)\big) - y_n = \varphi_n\big(g_n(y_n)\big) - f_n\big(g_n(y_n)\big)$$

it suffices by (17) to show that $g_n(y_n) \to y_0$. For this let $\epsilon > 0$ be given. The assumptions on f_n imply that $f_n(x) \to x$ for each x, and in particular for $x = y_0 \pm \epsilon$. Thus for all large n we have

$$f_n(y_0 - \epsilon) \le y_n \le f_n(y_0 + \epsilon);$$

these inequalities are equivalent to

$$y_0 - \epsilon \le g_n(y_n) \le y_0 + \epsilon$$

since $g_n = f_n^{-1}$ is an increasing function.

• (23) *holds*. As above, $g_n(y_n) \to y_0$. Since d_1 and d_2 are continuous, $\gamma_n(y_n) \to y_0$ as well. This implies that $x_n \to y_0$, since x_n lies between $g_n(y_n)$ and $\gamma_n(y_n)$. Consequently

$$\varphi_n'(x_n) = 1 + c_1'(x_n)b_n + c_2'(x_n)b_n^2 \to 1$$

since c_1' and c_2' are continuous. ∎

Theorem 4. *Suppose* (4) *holds with* $\Omega^{(2)}_{n,E}$ *defined by* (5) *and the remainder* $\rho^{(2)}_{n,E}(x) = O(1/n^{3/2})$ *locally uniformly in* x. *Then*

$$C_n(z) = z + \Omega^{(2)}_{n,C}(z) + O^*\Big(\frac{1}{n^{3/2}}\Big) \tag{24}$$

where

$$\begin{aligned}\Omega^{(2)}_{n,C}(z) &= \Big[\frac{\kappa_3 H_2(z)}{6}\Big]\frac{1}{\sqrt{n}}\\ &\quad + \Big[\frac{\kappa_4 H_3(z)}{24} + \frac{\kappa_3^2\big(-2H_3(z) - 1H_1(z)\big)}{36}\Big]\frac{1}{n}.\end{aligned} \tag{25}$$

$C_n^{(2)}(z) := z + \Omega^{(2)}_{n,C}(z)$ is called the second-order approximation to the Cornish-Fisher transformation.

Proof of Theorem 4. The asymptotic approximation (12) for N_n has the form

$$f_n(x) = x + c_1(x)b_n + c_2(x)b_n^2 + O^*(b_n^3)$$

where

$$\begin{aligned}&b_n = 1/\sqrt{n},\\ &f_n(x) = N_n(x),\\ &c_1(x) = -\kappa_3 H_2(x)/6,\\ &c_2(x) = -\big(\kappa_4 H_3(x)/24 + [\kappa_3^2(-4H_3(x) - 5H_1(x)]/36\big).\end{aligned}$$

Since c_1 and c_2 satisfy the smoothness requirements in Theorem 3, that result gives

$$\begin{aligned}C_n(z) &= N_n^{-1}(z) = f_n^{-1}(z)\\ &= z - c_1(z)b_n + \big(c_1(z)c_1'(z) - c_2(z)\big)b_n^2 + O^*(b_n^3)\\ &= z + \Omega^{(2)}_{n,C}(z).\end{aligned}$$ ∎

Remark If the distribution function F_n of X_n has an r^{th}-order Edgeworth expansion of the form

$$F_n(x) = \Phi(x) - \phi(x)\Omega^{(r)}_{n,E}(x) + O^*(1/n^{(r+1)/2}) \tag{26_E}$$

then N_n and C_n have r^{th}-order expansions of the form

$$N_n(x) = x - \Omega^{(r)}_{n,N}(x) + O^*(1/n^{(r+1)/2}) \tag{26_N}$$

$$C_n(x) = x + \Omega^{(r)}_{n,C}(x) + O^*(1/n^{(r+1)/2}) \tag{26_C}$$

which can be derived by the methods of this lecture. For simplicity of exposition, we have dealt just with the case $r = 2$. Each of the series $\Omega^{(r)}_{n,E}$, $\Omega^{(r)}_{n,N}$, and $\Omega^{(r)}_{n,C}$ has of the form

$$\Omega^{(r)}_{n,\alpha}(x) = P_1^\alpha(x)\frac{1}{\sqrt{n}} + P_2^\alpha(x)\frac{1}{n} + \cdots + P_r^\alpha(x)\frac{1}{n^{r/2}}$$

where the P_j^α's are polynomials. The choice of signs in (26) makes the leading term of each series the same.

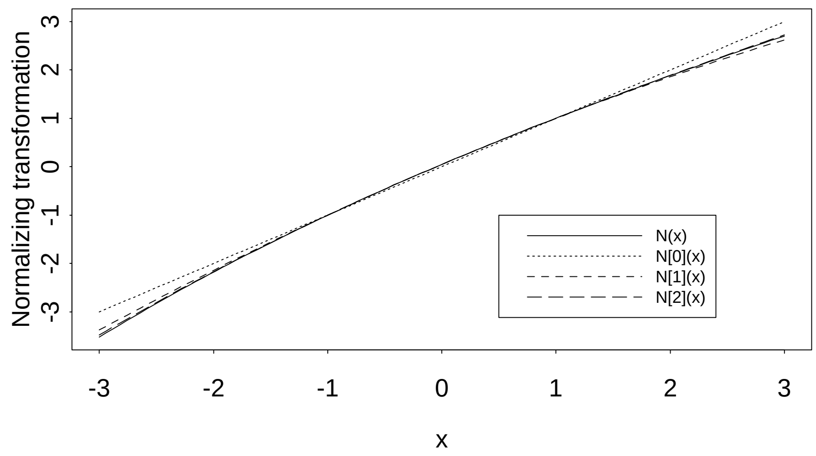
Example 1. Let Y_n have a Chisquare distribution with n degrees of freedom and consider the standardized variable

$$X_n = \frac{Y_n - E(Y_n)}{\mathrm{SD}(Y_n)} = \frac{Y_n - n}{\sqrt{2n}}.$$

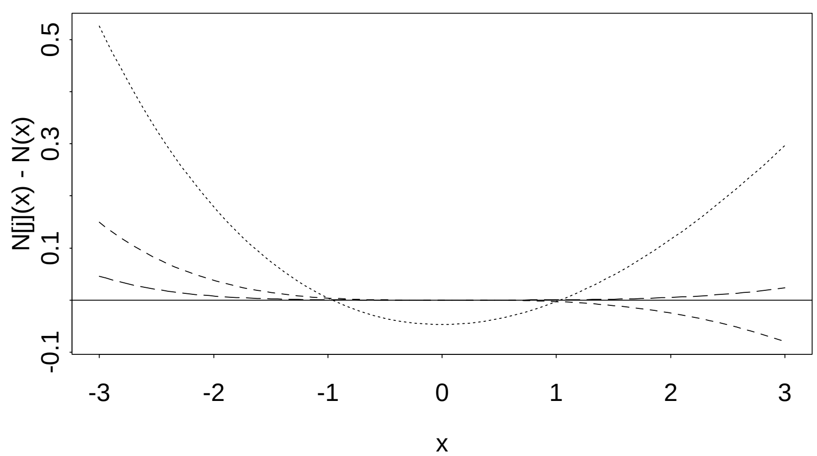
Since Y_n is the sum of n independent χ_1^2 random variables, Theorem 14.1 implies that the distribution function F_n of X_n has the Edgeworth expansion (4) with $\kappa_3 = \kappa_3(\chi_1^2)/2^{3/2} = \sqrt{8}$ and $\kappa_4 = \kappa_4(\chi_1^2)/2^{4/2} = 12$ (verify this!). Figure 1 exhibits certain features of the normalization transformation N_n for $n = 100$. The top panel plots $N_n(x)$ along with its 0^{th}, 1^{st}, and 2^{nd} order approximations $N_n^{(0)}(x)$, $N_n^{(1)}(x)$, and $N_n^{(2)}(x)$, for $-3 \le x \le 3$. The bottom panel plots the errors $N_n^{(j)}(x) - N_n(x)$ in these approximations, for $j = 1, 2, 3$. There are several things to note.

Figure 1

The normalization transformation for the standardized Chisquare(n) distribution and its 0:2nd order approximations for n = 100



Absolute errors in the N[0:2] approximations for n=100



- Since $(x, N_n(x)) = \big(F_n^{-1}(F_n(x)), \Phi^{-1}(F_n(x))\big)$, the graph of N_n is a plot of quantiles of Z against the corresponding quantiles of X_n. This Q-Q plot shows that the distribution of X_n has a thinner left-hand tail than that of Z (e.g., $N_n(-3) \approx -3.5$) but a thicker right-hand tail (e.g., $N_n(3) \approx 2.7$).
- $N_n^{(0)}(x) = x$ is simply the identity transformation, corresponding to the approximation $F_n(x) \approx \Phi(x)$. $N_n^{(0)}$ is not a very good approximation to N_n. $N_1^{(n)}$ and $N_2^{(n)}$ are moderately accurate over the range studied — $-3 \le x \le 3$.
- The error curves suggest (and it is indeed the case) that

$$\sup\{|N_n^{(j)}(x) - N_n(x)| : x \in \mathbb{R}\} = \infty$$

for each j and n. This is consistent with Theorem 2, which only asserts that $N_n^{(j)}(x) - N_n(x) = O(1/n^{(j+1)/2})$ locally uniformly in x as $n \to \infty$. And it is in sharp contrast to the Edgeworth expansion for F_n, for which we have (using a similar notation)

$$\sup\{|F_n^{(j)}(x) - F_n(x)| : x \in \mathbb{R}\} = O(1/n^{(j+1)/2}) = o(1).$$

Figure 2 is like Figure 1, but deals with the Cornish-Fisher transformation C_n (again for $n = 100$). Things to note here are:

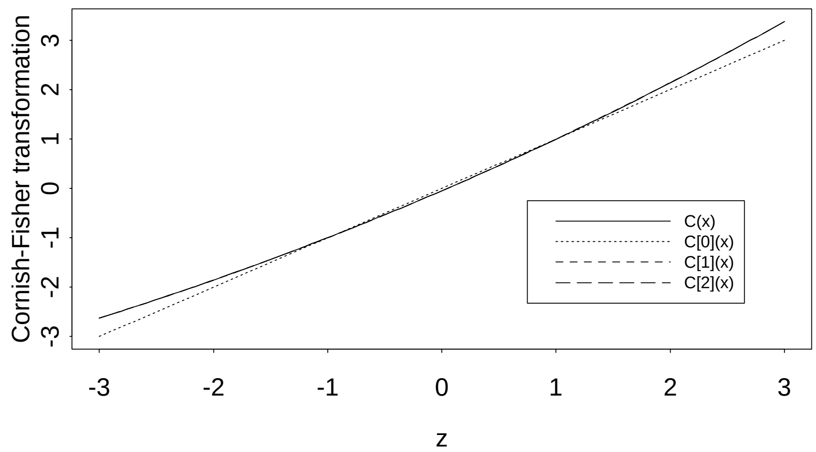
- Since $(z, C_n(z)) = \big(\Phi^{-1}(\Phi(z)), F_n^{-1}(\Phi(z))\big)$, the graph of C_n is a Q-Q plot of X_n against Z. Thus the graph of C_n in Figure 2 is essentially just the graph of N_n in Figure 1 with the x and z axes interchanged (the range of values plotted is different, and so are the scales).
- The first and second order approximations $C_n^{(1)}(z)$ and $C_n^{(2)}(z)$ to $C_n(z)$ are quite accurate over the range studied, i.e., $-3 \le z \le 3$.

Figure 3 plots

$$\epsilon_{N_n,j} := \log_{10}\Big(\sup\{\,\big|N_n^{(j)}(x) - N_n(x)\big| : -3 \le x \le 3\,\}\Big)$$
$$\epsilon_{C_n,j} := \log_{10}\Big(\sup\{\,\big|C_n^{(j)}(z) - C_n(z)\big| : -3 \le z \le 3\,\}\Big)$$

Figure 2

The Cornish-Fisher transformation for the standardized Chisquare(n) distribution and its 0:2nd order approximations for n = 100



Absolute errors in the C[1:2] approximations for n = 100

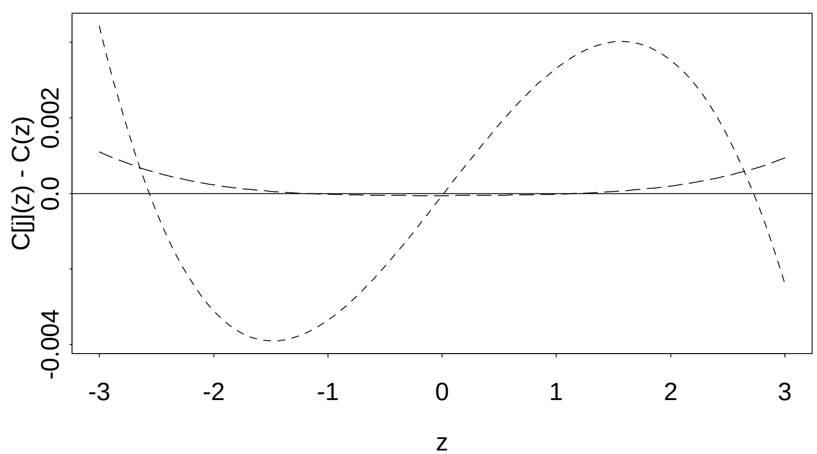
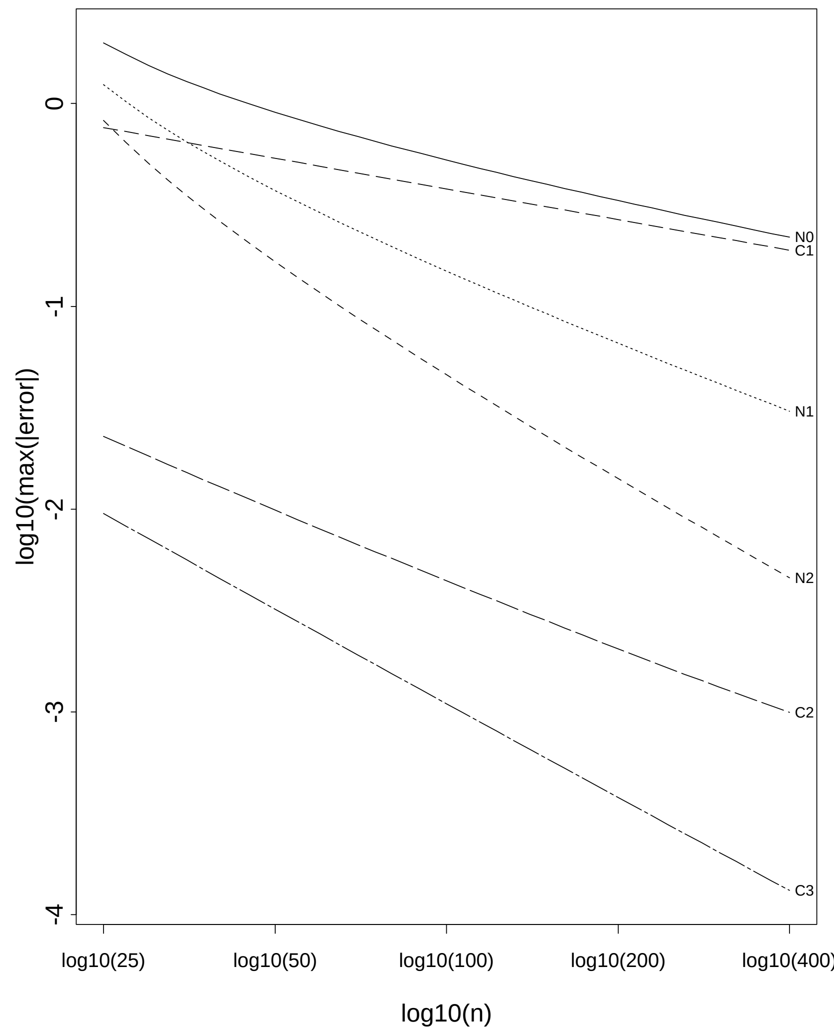


Figure 3

log10(max(|error|)) for the 0:2nd order normalization and Cornish-Fisher transformations for the standardized Chisquare(n) distribution, for n = 25 to 400



against $\log_{10} n$, for $j = 0, 1, 2$ and n running from 25 to 400. The plot shows among other things that (in this example) the 1st and 2nd order Cornish-Fisher transformations are about 25 times more accurate than the corresponding normalization transformations. Exercise 7 asks you to explain why. •

Exercise 1 (*Equivalent formulations of O^**). Let $f_1,\ f_2,\ \ldots$ be a infinite sequence of functions from $\mathbb{R}$ to $\mathbb{R}$ and let $\beta_1,\ \beta_2$, be an infinite sequence of positive real numbers. Show that the following are equivalent:

(O1) for each $x \in \mathbb{R}$, there exists an ϵ (which may depend on x) such that

$$\sup\{\, |f_n(\xi)| : |\xi - x| \leq \epsilon \,\} = O(\beta_n); \tag{27}$$

(O2) for every convergent sequence $(x_n)_{n=1}^{\infty}$ of real numbers,

$$f_n(x_n) = O(\beta_n); \tag{28}$$

(O3) for every bounded interval $[x_*, x^*]$

$$\sup\{\, |f_n(x)| : x \in [x_*, x^*]\,\} = O(\beta_n). \tag{29}$$

[Hint: If (29) fails for some $[x_*, x^*]$, then there exist strictly increasing indices n' and points $x'_n \in [x_*, x^*]$ such that $|f_{n'}(x_{n'})|/\beta_{n'} \to \infty$. Since $[x_*, x^*]$ is compact, the sequence $(x_{n'})$ has a convergent subsequence.] ◇

Exercise 2. Let $f_1,\ f_2,\ \ldots$ be a infinite sequence of functions from $\mathbb{R}$ to $\mathbb{R}$ and let $\beta_1,\ \beta_2$, be an infinite sequence of positive real numbers. Write $f_n(x) = O^{**}(\beta_n)$ to mean that $f_n(x)$ is $O(\beta_n)$ uniformly for $x \in \mathbb{R}$, i.e., $\sup\{\, |f_n(x)| : x \in \mathbb{R}\,\} = O(\beta_n)$. (a) Show that $f_n(x) = O^{**}(\beta_n)$ implies $f_n(x) = O^*(\beta_n)$. (b) Show by example that $f_n(x) = O^*(\beta_n)$ does not imply $f_n(x) = O^{**}(\beta_n)$. ◇

Exercise 3. Let R be the inverse to the cdf Φ of the standard normal distribution. Show that for $k \in \mathbb{N}$,

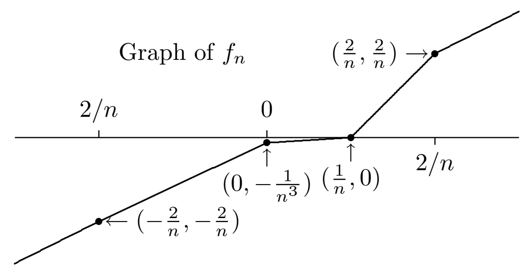
$$R^{(k)}\big(\Phi(x)\big) = \frac{q_k(x)}{\phi^k(x)}, \tag{30}$$

where the q_k's are polynomials satisfying the recursion relations

$$q_k(x) = (k-1)x q_{k-1}(x) + q'_{k-1}(x) \tag{31}$$

with $q_0(x) = x$. Write down q_1, q_2, q_3, and q_4 explicitly. ◇

Exercise 4. For $n \in \mathbb{N}$, let $f_n\colon \mathbb{R} \to \mathbb{R}$ be defined as follows: $f_n(x) = x$ for $|x| > 2/n$; over $[-2/n, 2/n]$, f_n is piecewise linear with vertices $(-2/n, -2/n)$, $(0, -1/n^3)$, $(1/n, 0)$, and $(2/n, 2/n)$:



Show that

$$f_n(x) = x + c_1(x)/n + c_2(x)/n^2 + O(1/n^3)$$

for each fixed x, with $c_1(x) = 0 = c_2(x)$. Let g_n be the inverse to f_n, and set $d_1(y) = 0 = d_2(y)$. Show that

$$g_n(y) = y + d_1(y)/n + d_2(y)/n^2 + O(1/n^3)$$

for each fixed $y \neq 0$, but not for $y = 0$. Why doesn't this contradict Theorem 3? ◇

Exercise 5. Formulate third order versions of Theorems 1 and 3 and use them to derive third order versions of Theorems 2 and 4. Recall that

$$\Omega^{(3)}_{n,E}(x) = \Omega^{(2)}_{n,E}(x) + \Big(\frac{\kappa_5 H_4(x)}{5!} + \frac{35\kappa_4\kappa_3 H_6(x)}{7!} + \frac{280\kappa_3^3 H_8(x)}{9!}\Big)\frac{1}{n^{3/2}}.$$

You may use Maple or the equivalent to do the algebra. ⋄

Exercise 6. Formulate and prove a version of Theorem 3 for functions f_n of the form

$$f_n(x) = c_0(x) + c_1(x)b_n + c_2(x)b_n^2 + O^*(b_n^3)$$

where c_0 is smooth and strictly increasing, and, moreover, c_0, c_1, and c_2 are allowed to depend on n. ⋄

Exercise 7. In the context of Example 1, let $c_N(x)$ be the coefficient of the $1/n$ term in $N_n^{(2)}(x)$, and, similarly, let $c_C(z)$ be the coefficient of the $1/n$ term in $C_n^{(2)}(z)$. Make a simultaneous plot of $c_N(y)$ and $c_C(y)$ against y, for $-3 \le y \le 3$. Use the plot to explain why $C_1^{(n)}$ is about 25 times as accurate as $N_1^{(n)}$ (see Figure 3). What is it about the distribution of X_1 — χ_1^2, standardized to mean 0 and variance 1 — that causes this phenomenon? ⋄

Exercise 8. Let $U_1, U_2, \ldots$ and $V_1, V_2, \ldots$ be independent standard exponential random variables, and set $S_n = \sum_{k=1}^n (V_k - U_k)$ for $n \ge 1$. (a) Show that S_n has density

$$f_{S_n}(y) = \sum\nolimits_{k=0}^{\nu}\Big[\binom{2\nu-k}{\nu}\frac{1}{2^{2\nu-k+1}}\Big]e^{-|y|}\frac{|y|^k}{k!} \tag{32}$$

for $-\infty < y < \infty$; here $\nu := n-1$. (b) Show that S_n has mean 0, variance $2n$, third cumulant 0, and fourth cumulant $12n$. [Hint: don't use (32).] (c) Write out $N_n^{(2)}(x)$ and $C_2^{(n)}(z)$ for the standardized variable $X_n := S_n/\sqrt{2n}$. (d) Plot $N_n^{(j)}(x) - N_n(x)$ versus $x \in [0,3]$ for $j = 0$, 1, 2 and $n = 25$ and 50. Similarly plot $C_n^{(j)}(z) - C_n(z)$ for the same j's and n's and $z \in [0,3]$. What do you conclude? ⋄

Problem 1. (*On the Δ-method*). For each $n \in \mathbb{N}$, let X_n be a random variable with a continuous strictly increasing distribution function F_n. Suppose that the inverse normalizing transformation $C_n = F_n^{-1}(\Phi)$ of X_n satisfies

$$C_n(z) = z + \frac{P_1(z)}{\sqrt{n}} + \frac{P_2(z)}{n} + O^*\Big(\frac{1}{n^{3/2}}\Big) \tag{33}$$

for some polynomials P_1 and P_2 which do not depend on n. Here "O^*" denotes "O", locally uniformly, and Φ is the distribution function of $Z \sim N(0,1)$. Equation (33) not only implies that X_n is asymptotically standard normal, but describes how the departure from normality diminishes with n.

(a) For each $n \in \mathbb{N}$, suppose that g_n is a continuous strictly increasing real-valued function of a real variable such that

$$g_n(x) = x + \frac{\gamma_1 x^2}{\sqrt{n}} + \frac{\gamma_2 x^3}{n} + O^*\Big(\frac{1}{n^{3/2}}\Big) \tag{34}$$

as $n \to \infty$. Put

$$X_n^* = g_n(X_n). \tag{35}$$

Show that the inverse normalizing transformation C_n^* of X_n^* satisfies

$$C_n^*(z) = z + \frac{P_1^*(z)}{\sqrt{n}} + \frac{P_2^*(z)}{n} + O^*\Big(\frac{1}{n^{3/2}}\Big) \tag{36}$$

for some polynomials P_1^* and P_2^*. Express the P_j^*'s in terms of the P_j's and γ_j's. ⋄

Part (a) yields the "Δ-method" with correction terms, as follows. Suppose that X_n above has the form

$$X_n = \frac{\sqrt{n}\,(Y_n - \mu)}{\sigma} \tag{37}$$

for some random variable Y_n and some numbers μ and $\sigma > 0$. Consider

$$X_n^* = \frac{\sqrt{n}\,\big(g(Y_n) - g(\mu)\big)}{\sigma g'(\mu)} \tag{38}$$

for a smooth function g. Since $X_n^* \approx X_n$, X_n^* is itself asymptotically standard normal. In fact, according to the next part, the inverse normalizing transformation C_n^* of X_n^* even satisfies (36).

(b) Let X_n^* be defined by (38), with g being four-times continuously differentiable with $g'(x) > 0$ for all x. Show that X_n^* can be written in the form $X_n^* = g_n(X_n)$ for functions g_n satisfying (34). Express the γ_j's of (34) in terms of g, μ, and σ^2. $\diamond$

The rest of the problem deals with the case where

$$Y_n = V_n/n \tag{39}$$

with $V_n \sim \chi_n^2$.

(c) Find μ and σ such that the X_n of (37) satisfies (33); write down the P_1 and P_2 of (33) explicitly. [Hint: use Theorem 4.] $\diamond$

(d) Let X_n^* be defined by (38) with g of the form $g(x) = x^c$ with $c > 0$. By parts (a)–(c), the inverse normalizing transformation C_n^* of X_n^* satisfies (36). Show that the polynomial P_1^* of (36) is constant if and only if $c = 1/3$, in which case $P_1^*(z) = -\sqrt{2/(9n)}$. Deduce that the inverse normalizing transformation C_n^{**} of

$$W_n := \frac{\sqrt[3]{\frac{V_n}{n}} - \left(1 - \frac{2}{9n}\right)}{\sqrt{\frac{2}{9n}}} \tag{40}$$

satisfies

$$C_n^{**}(z) = z + \frac{Q(z)}{n} + O^*\Big(\frac{1}{n^{3/2}}\Big) \tag{41}$$

for a polynomial Q; write out Q explicitly. $\diamond$

Part (d) says that the so-called ***Wilson-Hilferty transformation*** W_n of V_n is nearly standard normally distributed.

(e) Use Splus (or the equivalent) to study how close $C_n^{**}(z)$ is to z for small to moderate n and interesting values of z. Make some relevant plots and briefly discuss what conclusions you draw from them. $\square$

Problem 2. (*A high order version of Theorem 3*). This problem establishes a high order version of Theorem 3 (which itself is of second order). We begin with a preliminary result about the expansion of $(a_1x + a_2x^2 + \dots)^j$. To present this, we need to introduce some terminology and notation regarding additive partitions of an integer. Let k be a strictly positive integer. For $j = 1, 2, \dots$, let ${}_j\mathcal{P}_k$ be the collection of j-tuples $(i_1, i_2, \dots, i_j)$ of strictly positive integers such that

$$i_1 + i_2 + \cdots + i_j = k \quad \text{and} \quad i_1 \geq i_2 \geq \dots \geq i_j.$$

We call ${}_j\mathcal{P}_k$ the collection of ***additive partitions of k with j components***. Note that ${}_j\mathcal{P}_k = \emptyset$ if $j > k$. For $\pi \in {}_j\mathcal{P}_k$ let $\mu_i(\pi)$ be the number of times the integer i is a component of π, and note that $\sum_i \mu_i(\pi) = j$. For example, $\pi = (3,3,2,2,2,2,1,1,1) \in {}_9\mathcal{P}_{17}$ has $\mu_1(\pi) = 3$, $\mu_2(\pi) = 4$, $\mu_3(\pi) = 2$, and $\mu_4(\pi) = \mu_5(\pi) = \cdots = 0$; moreover $\sum_i \mu_i(\pi) = 3 + 4 + 2 = 9$.

(a) Let j and k be strictly positive integers with $j \leq k$. Show that the coefficient of x^k in the expansion of $(a_1x + a_2x^2 + a_3x^3 + \cdots)^j$ is

$$C\big(k, j, (a_1, a_2, a_3, \dots)\big) := j! \sum\nolimits_{\pi \in {}_j\mathcal{P}_k} \left(\prod\nolimits_i \frac{a_i^{\mu_i(\pi)}}{(\mu_i(\pi))!} \right). \qquad (42)\ \diamond$$

The goal now is to establish

Theorem 5. *Suppose $f_1, f_2, \dots$ is an infinite sequence of strictly increasing continuous functions having an asymptotic expansion of the form*

$$f_n(x) = x + c_1(x)b_n + c_2(x)b_n^2 + \cdots + c_\ell(x)b_n^\ell + O^*(b_n^{\ell+1}) \qquad (43)$$

where ℓ is a positive integer, c_k is $(\ell - k + 1)$-times continuously differentiable for $1 \leq k \leq \ell$, and $b_n = o(1)$. Then the inverse functions $g_n = f_n^{-1}$ have the asymptotic expansion

$$g_n(y) = y + d_1(y)b_n + d_2(y)b_n^2 + \cdots + d_\ell(y)b_n^\ell + O^*(b_n^{\ell+1}) \qquad (44)$$

where the functions $d_1(y), \dots, d_\ell(y)$ are defined recursively by the equations

$$\begin{aligned} d_k(y) &= -c_k(y) \\ &\quad - \Big(\sum_{i=1}^{k-1} \sum_{j=1}^{k-i} \frac{c_i^{(j)}(y)}{j!} C\big(k-i, j, (d_1(y), d_2(y), \dots)\big) \Big) \end{aligned} \qquad (45_k)$$

for $k = 1, \dots, \ell$, with C given by (42).

(b) Show that the equations (45) really are recursive, i.e., that the right-hand side of (45_k) only references $d_1(y), \dots, d_{k-1}(y)$. $\diamond$

The proof of Theorem 5 is for the most part a straight-forward generalization of the proof of Theorem 3, so I won't ask you to write it out. There is one issue though that needs some thought, namely, the ℓ^{th}-order analogue of (22). This is addressed by the next part.

(c) For arbitrary x and y set

$$\begin{aligned} \varphi_n(x) &= x + c_1(x)b_n + \cdots + c_\ell(x)b_n^\ell \\ \gamma_n(y) &= y + \Delta_n(y) \quad \text{with} \quad \Delta_n(y) = d_1(y)b_n + \cdots + d_\ell(y)b_n^\ell. \end{aligned}$$

Show that if $y_n \to y_0$ in $\mathbb{R}$, then $\varphi_n\big(\gamma_n(y_n)\big) = y_n + O(b_n^{\ell+1})$. [Hint: use part (a) to generalize the proof of (22); justify each step of your argument.] $\diamond$

The rest of the problem deals with some properties of the d_k's.

(d) By (19), $d_1 = -c_1$ and $d_2 = c_1c_1' - c_2$. Show that

$$d_3 = -c_1(c_1')^2 - \tfrac{1}{2}c_1^2c_1'' + c_1'c_2 + c_1c_2' - c_3, \qquad (46)$$

$$\begin{aligned} d_4 = c_1(c_1')^3 &+ \tfrac{3}{2}c_1^2c_1'c_1'' + \tfrac{1}{6}c_1^3c_1''' \\ &- (c_1')^2c_2 - 2c_1c_1'c_2' - \tfrac{1}{2}c_1^2c_2'' - c_1c_1''c_2 \\ &+ c_2c_2' + c_1c_3' + c_1'c_3 - c_4. \end{aligned} \qquad (47)$$

You may use Maple or the equivalent to do the algebra. $\diamond$

(e) Show that for each k, d_k is a linear combination of terms having the following structure. For j running from 1 to k, take each additive partition $\pi = (i_1, \ldots, i_j)$ of k into j components and consider the product $c_\pi := c_{i_1} c_{i_2} \cdots c_{i_j}$. d_k has one term for each distinct way of applying $j-1$ derivative signs $'$ to the j factors of c_π. [Hint: To understand the claim, look at (47). To prove the claim, use (45) and induction on k.]

(f) Show that d_k is $(\ell - k + 1)$-times continuously differentiable. [Hint: Use (45) and induction on k; alternatively, use the result of part (e).] $\diamond$

(g) Show that if $c_1(x) = \cdots = c_\ell(x) = 0$ for some $x \in \mathbb{R}$, then $d_1(y) = \cdots = d_\ell(y) = 0$ for $y = x$. $\square$

Problem 3. (*On the quantiles of the normal distribution*). Let Z be a standard normal random variable. For $0 < p < 1$, let z_p be the p^{th} quantile of the distribution of Z, i.e., the number such that

$$\Phi(z_p) := P[Z \le z_p] = p. \tag{48}$$

This problem develops an asymptotic expansion for z_p as $p \uparrow 1$. Actually, the asymptotic expansion is for the quantity x_p such that

$$z_p = \sqrt{v_p(1 - x_p)} \tag{49}$$

where $v_p := 2\log(1/(1-p))$. It would be easy to obtain an expansion for z_p from the expansion for x_p, but there is no advantage to doing so. Suppose throughout that $p \ge \Phi(1)$, so $z_p \ge 1$. For notational convenience, set $u_p = 1/v_p$.

(a) Show that x_p is the root x of the equation $f_p(x) = y_p$, where

$$f_p(x) := x - u_p \log(1-x) + 2u_p \big[\log R\big(u_p/(1-x)\big) - \log R(u_p)\big]$$

and

$$y_p := \xi_p u_p - 2u_p \log R(u_p)$$

with

$$\xi_p := \log(2\pi v_p).$$

Here R is the function defined implicitly by

$$R(1/z^2) = \frac{1 - \Phi(z)}{\phi(z)/z}$$

with $\phi = \Phi'$. $\diamond$

(b) Show that $f_p(0) = 0$ and that $f_p'(x) \ge 1$. [Hint. $f_p'(x)$ is a function, say F, just of the variable $a = u_p/(1-x)$. If you have trouble showing analytically that $F(a) \ge 1$, draw a convincing graph.] $\diamond$

(c) Show that $\log R$ has an asymptotic expansion of the form

$$(\log R)(\zeta) \asymp \sum_{k=1}^{\infty} \rho_k \zeta^k$$

as $\zeta \downarrow 0$, with $\rho_1 = -1$, $\rho_2 = 5/2$, $\rho_3 = -37/3$, and $\rho_4 = 353/4$. [Hint: Use Theorem 1.1.] $\diamond$

(d) Show that as $p \uparrow 1$, $f_p(x)$ has the asymptotic expansion

$$f_p(x) \asymp x + \sum_{k=1}^{\infty} c_k(x) u_p^k \tag{50}$$

where

$$c_k(x) = \begin{cases} -\log(1-x), & \text{if } k = 1, \\ 2\rho_{k-1}\big[1/(1-x)^{k-1} - 1\big], & \text{if } k \geq 2. \end{cases}$$

Verify that for each ℓ the error incurred in terminating this expansion with the term for $k = \ell$ is $O(u_p^{\ell+1})$, uniformly for $x \in [0, 1/2]$. $\diamond$

Now let $g_p = f_p^{-1}$ be the inverse to f_p. It follows from Theorem 5 that g_p has the asymptotic expansion

$$g_p(y) \asymp y + \sum_{k=1}^{\infty} d_k(y) u_p^k \tag{51}$$

as $p \uparrow 1$, where the functions d_1, d_2, … are defined in terms of c_1, c_2, … by (45); moreover for each ℓ the error incurred in terminating this expansion with the term for $k = \ell$ is $O(u_p^{\ell+1})$, uniformly for $y \in [0, 1/2]$. (Theorem 5 was formulated for functions f defined on all of $\mathbb{R}$, but the argument can be easily modified to handle functions defined on a bounded interval; you do not have to give this argument.)

To get the desired asymptotic expansion of x_p, operate formally, as follows. In (51), expand $d_1(y)$, $d_2(y)$, … in power series about 0, substitute $\xi u_p - 2\rho_1 u_p^2 - 2\rho_2 u_p^3 - \cdots$ for y, and rearrange the terms into an expression of the form

$$t_1(\xi) u_p + t_2(\xi) u_p^2 + t_3(\xi) u_p^3 + \cdots,$$

where $t_1(\xi)$, $t_2(\xi)$, … depend just on ξ.

(e) Show that

$$\begin{aligned} t_1(\xi) &= \xi, \\ t_2(\xi) &= 2 - \xi, \\ t_3(\xi) &= (-14 + 6\xi - \xi^2)/2, \\ t_4(\xi) &= (214 - 102\xi + 21\xi^2 - 2\xi^3)/6, \\ t_5(\xi) &= (-2978 + 1488\xi - 348\xi^2 + 46\xi^3 - 3\xi^4)/12. \end{aligned}$$

You may use Maple or the equivalent to do the algebra. $\diamond$

(f) Show that for all $k \geq 2$, $t_k(\xi)$ is a polynomial in ξ of degree $k-1$. $\diamond$

(g) Show that

$$x_p \asymp \sum_{k=1}^{\infty} t_k(\xi_p) u_p^k, \tag{52}$$

in the sense that for each $\ell \geq 1$ the error committed by terminating the expansion with the term for $k = \ell$ is $O(\xi_p^\ell u_p^{l+1})$ as $p \uparrow 1$. [Hint: Parts (f) and (g) of the preceding problem are helpful.] $\diamond$

(h) (The proof of the pudding is in the eating.) For $\ell = 0, 1, \ldots$, put $\hat{z}_{p,\ell} = \sqrt{v_p(1 - \hat{x}_{p,\ell})}$, where $\hat{x}_{p,\ell}$ is the sum of the first ℓ terms of the expansion (52) (take $\hat{x}_{p,0} = 0$.) Let $\epsilon_{p,\ell} = \hat{z}_{p,\ell}/z_p - 1$ be the relative error in $\hat{z}_{p,\ell}$ as an approximation to z_p. Use Splus or the equivalent to draw some plots that exhibit $\epsilon_{p,\ell}$ in an informative manner. $\square$