Amanda Fernández Fontelo

November 14, 2019

1 Outline paper fattering-thinning operator

1. Introduction.
2. The model for misreporting counts:
   (a) Introduction to the fattering-thinning operator.
   (b) Introduction to the misreporting model accounting for two different latent process structures: a. the classical INAR(1) with Poisson(λ)-innovations, and b. an INAR(1) model with 2nd-Hermite(a_1,a_2)-innovations. The latter means that the latent process X_n can be slight overdispersed. Also, considering both independence and dependence between the misreporting states.
   Could also be interesting to briefly describe, in some way how, is an INAR(1) model with the fattering-thinning operator (marginal distributions, ways of parameter estimation, etc).
3. Simulation study.
4. Applications based on ADHD and autism cases in children.
5. Discussion.

2 The fattering-thinning operator

Let X_n be a latent process following an INAR(1) structure such that: $X_n = \alpha \circ X_{n-1} + Z_n$, where $\mathrm{E}(X_n) = \mu_X$ and $\mathrm{Var}(X_n) = \sigma^2_X$ are the expectation and variance of X_n, respectively. Assume, for now, that $Z_n \sim \text{Poisson}(\lambda)$. However, other more appropriate structures for the underlying process can be considered depending on the application (e.g., overdispersion in the underlying process through second-order Hermite distributed innovations, or temporally uncorrelated processes through the Poisson or second-order Hermite models). Let Y_n be an observed and potentially over- or under-reporting process such that:

$$Y_n = \begin{cases} X_n & 1-\omega \\ \theta \Diamond X_n & \omega, \end{cases} \tag{1}$$

where $\theta \Diamond X_n$ is the fattering-thinning operator in the sense that:

$$\theta \Diamond X_n | X_n = x_n = \sum_{j=1}^{x_n} W_j, \tag{2}$$

where W_j is an i.i.d random variable defining by the following probability mass function (pmf):

$$\mathrm{P}(W_j = k) = \begin{cases} 0 & 1-\phi_1-\phi_2 \\ 1 & \phi_1 \\ 2 & \phi_2, \end{cases} \tag{3}$$

where $\theta = (\phi_1, \phi_2)$. Notice that when $\phi_2 = 1$, the process is not over-reported neither under-reported. That is, the observed process is the actual process. Notice also that a more restricted version of (2) results when $W_j \sim \text{Bernoulli}(2, \phi)$. Although the distribution in (3) is the most straightforward choice to allow for over-reporting, other distributions with not compact support can be considered, such as Poisson, Geometric, etc.
The operator in (2) follows a 2nd-Hermite distribution with parameters $\mu_X\phi_1$ and $\mu_X\phi_2$. This can be easily demonstrated by taking its probability generating function (pgf). That is:

$$G_X(s) = \mathrm{e}^{\mu_X(s-1)}, \tag{4}$$

$$G_W(s) = (1 - \phi_1 - \phi_2) + \phi_1 s + \phi_2 s^2, \tag{5}$$

$$G_X\left(G_W(s)\right) = \mathrm{e}^{\mu_X\left((1-\phi_1-\phi_2)+\phi_1 s+\phi_2 s^2-1\right)} = \mathrm{e}^{\mu_X\left(\phi_1(s-1)+\phi_2(s^2-1)\right)}, \tag{6}$$

which is the pgf of a second-order Hermite with parameters $\mu_X\phi_1$ and $\mu_X\phi_2$. The expectation and variance of this operator are: $\mathrm{E} = (\theta \Diamond X_n) = \mu_X\left(\phi_1 + 2\phi_2\right)$ and $\mathrm{Var} = (\theta \Diamond X_n) = \mu_X\left(\phi_1 + 4\phi_2\right)$.

2.1 Models properties

The marginal distribution of the observed process Y_n is the following mixture of a Poisson and a Hermite distributions:

$$Y_n = \begin{cases} \text{Poisson}(\mu_X) & 1-\omega, \\ \text{Hermite}\left(\mu_X\phi_1, \mu_X\phi_2\right) & \omega. \end{cases} \tag{7}$$

However, in the case that the innovations of the INAR(1)-latent process are second-order Hermite, then the distribution in (7) is a mixture of two second-order Hermite components. More straightforward marginal distributions of Y_n result when the latent process X_n follows a classical Poisson model, or even a second-order Hermite model.
The expectation and variance of this observed process Y_n are:

$$\begin{aligned}
\mathrm{E}(Y_n) &= (1-\omega)\mu_X + \omega\mu_X\left(\phi_2 + 2(1-\phi_1-\phi_2)\right) = \mu_X(1-\omega\left(1-(2(1-\phi_1)-\phi_2)\right)). \\
\mathrm{E}(Y_n^2) &= (1-\omega)(\sigma_X^2+\mu_X^2) + \omega\left(\mu_X(4(1-\phi_1)-3\phi_2) + \mu_X^2(2(1-\phi_1)-\phi_2)^2\right) \\
&= \mu_X\left(1-\omega\left(1-(4(1-\phi_1)-\phi_2)\right)\right) + \mu_X^2\left(1-\omega\left(1-(2(1-\phi_1)-\phi_2)^2\right)\right),
\end{aligned}$$

since $\mu_X = \sigma_X^2$, and

$$\begin{aligned}
\mathrm{Var}(Y_n) &= \mu_X\left(1-\omega\left(1-\left(4\left(1-\phi_1\right)-\phi_2\right)\right)\right) + \mu_X^2\left(1-\omega\left(1-\left(2\left(1-\phi_1\right)-\phi_2\right)^2\right)\right) \\
&\quad - \mu_X^2\left(1-\omega\left(1-\left(2\left(1-\phi_1\right)-\phi_2\right)\right)\right)^2 \\
&= \mu_X\left(1-\omega\left(1-\left(4\left(1-\phi_1\right)-\phi_2\right)\right)\right) + \mu_X^2\omega(1-\omega)\left(1-\left(2\left(1-\phi_1\right)-\phi_2\right)\right)^2.
\end{aligned}$$

Let $\mathbf{1}_n$ be an indicator of the state of misreporting (over- or under-reporting) in the sense that $\mathbf{1}_n \sim$ Bernoulli(ω). Assume that the misreporting states are independent over time.

The auto-covariance function (ACV) of Y_n can be computed as follows:

$$\mathrm{Cov}\left(Y_n, Y_{n+k}\right) = \mathrm{E}\left(Y_n, Y_{n+k}\right) - \mathrm{E}(Y_n)\mathrm{E}(Y_{n+1}). \tag{8}$$

Assume that $\mathbf{1}_n \sim$ Bernoulli(ω) independent of X_n. Additionally, assume independence betwen the misreporting states. Hence:

$$\begin{aligned}
\mathrm{E}\left(Y_n, Y_{n+k}\right) &= \mathrm{E}\left(X_n(1-\mathbf{1}_n), X_{n+k}(1-\mathbf{1}_{n+k})\right) + \mathrm{E}\left(X_n(1-\mathbf{1}_n), \theta\Diamond X_{n+k}\mathbf{1}_{n+k}\right) \\
&\quad + \mathrm{E}\left(\theta\Diamond X_n\mathbf{1}_n, X_{n+k}(1-\mathbf{1}_{n+k})\right) + \mathrm{E}\left(\theta\Diamond X_n\mathbf{1}_n, \theta\Diamond X_{n+k}\mathbf{1}_{n+k}\right),
\end{aligned} \tag{9}$$

where

$$\begin{aligned}
&\mathrm{E}\left(X_n(1-\mathbf{1}_n), X_{n+k}(1-\mathbf{1}_{n+k})\right) = (1-\omega)^2\mathrm{E}(X_n, X_{n+k}),\\
&\mathrm{E}\left(X_n(1-\mathbf{1}_n), \theta\Diamond X_{n+k}\mathbf{1}_{n+k}\right) = (1-\omega)\omega(2(1-\phi_1)-\phi_2)\mathrm{E}(X_n, X_{n+k}),\\
&\mathrm{E}\left(\theta\Diamond X_n\mathbf{1}_n, \theta\Diamond X_{n+k}\mathbf{1}_{n+k}\right) = \omega^2(2(1-\phi_1)-\phi_2)^2\mathrm{E}(X_n, X_{n+k}).
\end{aligned}$$

Accordingly,

$$\begin{aligned}
\mathrm{E}\left(Y_n, Y_{n+k}\right) &= (1-\omega)^2\mathrm{E}(X_n, X_{n+k}) + 2(1-\omega)\omega(2(1-\phi_1)-\phi_2)\mathrm{E}(X_n, X_{n+k})\\
&\quad + \omega^2(2(1-\phi_1)-\phi_2)^2\mathrm{E}(X_n, X_{n+k})\\
&= \mathrm{E}(X_n, X_{n+k})\left((1-\omega)^2 + 2(1-\omega)\omega(2(1-\phi_1)-\phi_2) + \omega^2(2(1-\phi_1)-\phi_2)^2\right)\\
&= \mathrm{E}(X_n, X_{n+k})\left((1-\omega) + \omega(2(1-\phi_1)-\phi_2)\right)^2\\
&= \mathrm{E}(X_n, X_{n+k})(1-\omega(1-(2(1-\phi_1)-\phi_2)))^2.
\end{aligned}$$

Finally,

$$\begin{aligned}
\mathrm{Cov}\left(Y_n, Y_{n+k}\right) &= (\sigma_X^2\alpha^k + \mu_X^2)(1-\omega(1-(2(1-\phi_1)-\phi_2)))^2 - \mu_X^2(1-\omega\left(1-(2(1-\phi_1)-\phi_2)\right))^2\\
&= \mu_X\alpha^k(1-\omega(1-(2(1-\phi_1)-\phi_2)))^2
\end{aligned}$$

where $\mathrm{E}(X_n) = \mu_X = \sigma_X^2 = \mathrm{Var}(X_n) =$ and $\mathrm{E}(X_n, X_{n+k}) = (\sigma_X^2\alpha^k + \mu_X^2)$.

The auto-correlation function (ACF) of the observed process takes then the following expression:

$$\mathrm{Cor}\left(Y_n, Y_{n+k}\right) = \frac{\alpha^k(1-\omega(1-(2(1-\phi_1)-\phi_2)))^2}{\left(1-\omega\left(1-(4\left(1-\phi_1\right)-\phi_2)\right)\right) + \mu_X\omega(1-\omega)\left(1-(2\left(1-\phi_1\right)-\phi_2)\right)^2} = \alpha^k\mathrm{c}(\alpha,\lambda,\omega,\phi_1,\phi_2).$$

The above computations can be extended to the case where the misreporting states are correlated through a two-state Markov chain. Suppose this new model, which assumes misreporting and a dependence structure between these misreporting states, is R_n.
As a result, the expectation and variance of this new process, that is, $\mathrm{E}(R_n)$ and $\mathrm{Var}(R_n)$ remain equal to those presented above. The latter is because the marginal distribution of R_n is the same than of Y_n, that is, the mixture of a Poisson distribution and 2nd-Hermite distribution (7).

However, this is not true for the auto-covariance and auto-correlation function of R_n, which should be more complex. In this sense, the covariance of the process R_n can be computed as follows.

Recall that $\mathbf{1}_n$ is an indicator of the state of misreporting (over- or under-reporting) in the sense that $\mathbf{1}_n \sim$ Bernoulli(ω). Suppose now that there is a dependence structure among these states through a binary Markov chain. From Fernández *et al.* (submitted), the transition probability $\mathbf{P}$ is:

$$\mathbf{P} = \begin{bmatrix} 1-p_{01} & p_{01} \\ p_{01}\frac{1-\omega}{\omega} & 1-p_{01}\frac{1-\omega}{\omega} \end{bmatrix}.$$

As proved in Fernández *et al.* (submitted), the $\mathbf{P}^k$ transition matrix can be written in terms of p_{01}^k. It is also worthy to consider that the parameter p_{01} can be written in terms of the second eigenvalue of $\mathbf{P}$, that is, $p_{01} = \omega(1-\lambda_2)$, where λ_2 is the second value of $\mathbf{P}$.

Assume that processes $\mathbf{1}_n$ and R_n are mutually independent. Similarly to expression (8) and (9), then:

$$\begin{aligned}
\mathrm{E}\left(X_n(1-\mathbf{1}_n), X_{n+k}(1-\mathbf{1}_n)\right) &= \mathrm{E}\left(X_n, X_{n+k}\right)\mathrm{P}(\mathbf{1}_n = 0, \mathbf{1}_{n+k} = 0) = \mathrm{E}\left(X_n, X_{n+k}\right)(1-\omega)(1-\omega(1-\lambda_2^k)),\\
\mathrm{E}\left(X_n(1-\mathbf{1}_n), \theta\Diamond X_{n+k}\mathbf{1}_n\right) &= \mathrm{E}\left(X_n, X_{n+k}\right)(2(1-\phi_1)-\phi_2)\mathrm{P}(\mathbf{1}_n = 0, \mathbf{1}_{n+k} = 1)\\
&= \mathrm{E}\left(X_n, X_{n+k}\right)(2(1-\phi_1)-\phi_2)\omega(1-\omega)(1-\lambda_2^k),\\
\mathrm{E}\left(\theta\Diamond X_n\mathbf{1}_n, \theta\Diamond X_{n+k}\mathbf{1}_n\right) &= \mathrm{E}\left(X_n, X_{n+k}\right)(2(1-\phi_1)-\phi_2)^2\mathrm{P}(X_n = 1, X_{n+k} = 1)\\
&= \mathrm{E}\left(X_n, X_{n+k}\right)(2(1-\phi_1)-\phi_2)^2\omega(1-(1-\omega)(1-\lambda_2^k)).
\end{aligned}$$

From the computations above:

$$\mathrm{E}(R_n, R_{n+k}) = \mathrm{E}(X_n, X_{n+k}) \left(\left(1 - \omega\left(1 - (2(1-\phi_1) - \phi_2)^2\right)\right) - \omega(1-\omega)(1 - (2(1-\phi_1) - \phi_2))^2(1-\lambda_2^k)\right) =$$

$$= (\alpha^k\sigma_X^2 + \mu_X^2)\left(\left(1 - \omega\left(1 - (2(1-\phi_1) - \phi_2)^2\right)\right) - \omega(1-\omega)(1 - (2(1-\phi_1) - \phi_2))^2(1-\lambda_2^k)\right)^2,$$

and the ACV function takes the following expression:

$$\mathrm{Cov}(R_n, R_{n+k}) = \mathrm{E}(R_n, R_{n+k}) - \mathrm{E}(R_n)\mathrm{E}(R_{n+k}) =$$

$$= (\alpha^k\sigma_X^2 + \mu_X^2)\left(\left(1 - \omega\left(1 - (2(1-\phi_1) - \phi_2)^2\right)\right) - \omega(1-\omega)(1 - (2(1-\phi_1) - \phi_2))^2(1-\lambda_2^k)\right)$$

$$- \mu_X^2\left(1 - \omega\left(1 - (2(1-\phi_1) - \phi_2)\right)\right)^2 =$$

$$= (\alpha^k\sigma_X^2 + \mu_X^2)\left(1 - \omega\left(1 - (2(1-\phi_1) - \phi_2)\right)\right)^2 + (\alpha^k\sigma_X^2 + \mu_X^2)\left(\lambda_2^k\omega(1-\omega)(1 - (2(1-\phi_1) - \phi_2))^2\right)$$

$$- \mu_X^2\left(1 - \omega\left(1 - (2(1-\phi_1) - \phi_2)\right)\right)^2 =$$

$$= \alpha^k\sigma_X^2\left(1 - \omega\left(1 - (2(1-\phi_1) - \phi_2)\right)\right)^2 + \mu_X^2\lambda_2^k\omega(1-\omega)(1 - (2(1-\phi_1) - \phi_2))^2 +$$

$$+ \sigma_X^2(\alpha\lambda)^k\omega(1-\omega)(1 - (2(1-\phi_1) - \phi_2))^2.$$

Finally, the ACF can be written as follows:

$$\mathrm{Cor}(R_n, R_{n+k}) = \frac{\alpha^k\left(1 - \omega\left(1 - (2(1-\phi_1) - \phi_2)\right)\right)^2 + \mu_X\lambda_2^k\omega(1-\omega)(1 - (2(1-\phi_1) - \phi_2))^2 + (\alpha\lambda)^k\omega(1-\omega)(1 - (2(1-\phi_1) - \phi_2))^2}{\left(1 - \omega\left(1 - (4(1-\phi_1) - \phi_2)\right)\right) + \mu_X\omega(1-\omega)\left(1 - (2(1-\phi_1) - \phi_2)\right)^2} \tag{10}$$

2.2 Parameters estimation

2.2.1 Moment-based method

2.2.2 Likelihood-based method

The parameters of the model can be estimated through the likelihood function. To do so, the forward algorithm can be used since the direct computation of the likelihood is not tractable. More details on the computations of the expression of the likelihood function of model based on the forward algorithm can be found in Fernández-Fontelo *et al.* (2016, 2019). In the scenario considered in this work, the forward probabilities can be computed following the expression:

$$\gamma_n\left(\boldsymbol{y}_{1:n}, x_n\right) = \sum_{x_{n-1}} \mathrm{P}\left(Y_n = y_n | X_n = x_n\right) \mathrm{P}\left(X_n = x_n | X_{n-1} = x_{n-1}\right) \gamma_{n-1}\left(\boldsymbol{y}_{1:n-1}, x_{n-1}\right) \tag{11}$$

where the emission probabilities take the following expression:

$$\mathrm{P}(Y_n = y_n | X_n = x_n) = \begin{cases} 0 & \text{if} \quad y_n < x_n < \frac{y_n}{2}, \\ (1-\omega) + \omega\frac{x_n!}{n_0!n_1!n_2!}(1-\phi_1-\phi_2)^{n_0}\phi_1^{n_1}\phi_2^{n_2} & \text{if} \quad y_n = x_n, \\ \omega\frac{x_n!}{n_0!n_1!n_2!}(1-\phi_1-\phi_2)^{n_0}\phi_1^{n_1}\phi_2^{n_2} & \text{if} \quad \frac{y_n}{2} \leq x_n < y_n, \\ 1 & \text{if} \quad y_n = 0, \end{cases} \tag{12}$$

being n_0, n_1 and n_2 the number of 0, 1 and 2, respectively, in a sequence of length x_n (e.g., $\{W_1 = w_1, W_2 = w_2, \ldots W_{x_n} = w_{x_n}\}$) restricted to $\sum_{i=1}^{x_n} w_i = y_n$. Given the observed value of y_n and a potential value of x_n, the number of possible sequences of 0, 1 and 2 keeping the above mentioned conditions is likely greater than one. In that case, the sum of probabilities of each possible sequence will be the emission probability of $Y_n = y_n | X_n = x_n$. Notice that in the case that $W_i \sim \text{Binomial}(2, \phi)$, the emission probabilities are as follows:

$$\mathrm{P}(Y_n = y_n | X_n = x_n) = \begin{cases} 0 & \text{if} \quad y_n < x_n < \frac{y_n}{2}, \\ (1-\omega) + \omega\binom{2x_n}{y_n}\phi^{y_n}(1-\phi)^{2x_n - y_n} & \text{if} \quad y_n = x_n, \\ \omega\binom{2x_n}{y_n}\phi^{y_n}(1-\phi)^{2x_n - y_n} & \text{if} \quad \frac{y_n}{2} \leq x_n < y_n, \\ 1 & \text{if} \quad y_n = 0. \end{cases} \tag{13}$$

The transition probabilities remain the same to those proposed in the under-reporting scenarios since the underlying process still follows the same model. However, other interesting structures for the latent process can be considered such as an INAR(1) process with second-order Hermite distributed innovations or even a temporally uncorrelated model (e.g., Poisson or second-order Hermite models, etc).
Finally, the likelihood function of the process $\{Y_n\}$ can be computed recursively through:

$$\mathrm{P}(\boldsymbol{Y}_{1:N} = \boldsymbol{y}_{1:N}) = \mathrm{P}\left(Y_1 = y_1, Y_2 = y_2, \ldots, Y_N = y_N\right) = \sum_{x_N = \frac{y_N}{2}}^{y_N} \gamma_N\left(\boldsymbol{y}_{1:N}, x_N\right), \tag{14}$$

taking $\mathrm{P}(X_1 = x_1) = \text{Poisson}(\frac{\lambda}{1-\alpha})$, or Hermite() in case the innovations of the INAR(1) process are second-order Hermite distributed.

The results above can be extended to the case of dependence between the states of misreporting. (TO DO)

3 Simulation study

4 Application

5 Discussion

Appendix

INAR(1) model based on the fattering-thinning operator

Suppose a version of the INAR(1) model with the following structure:

$$X_n = \theta \Diamond X_{n-1} + Z_n \tag{15}$$

where $X_n \sim \text{Poisson}(\lambda)$, and $\theta \Diamond X_{n-1}$ the fattering-thinning operator (2) and (3). Taking the pgf of X_n (4) and $\theta \Diamond X_{n-1}$ (6), the pgf function of Z_n takes the following expression:

$$\mathrm{G}_Z(s) = \frac{\mathrm{e}^{\lambda(s-1)}}{\mathrm{e}^{\lambda(\phi_2(s-1)+(1-\phi_1-\phi_2)(s^2-1))}} = \mathrm{e}^{\lambda\left((1-\phi_2)(s-1)+(\phi_1+\phi_2-1)(s^2-1)\right)}. \tag{16}$$

Expression (16) reminds the pgf of a 2nd-Hermite distribution. However, according to Kemp and Kemp (1965), both parameters of the 2nd-Hermite should be positive. In this case, $\phi_1 + \phi_2 - 1 < 0$ and, hence, the expression (16) is not a pgf. The latter means that the marginal distribution of the process (15) cannot be Poisson.

On the other hand, suppose now that $X_n \sim \text{Hermite}(a_1, a_2)$, then

$$\begin{aligned}
\mathrm{G}_X(\mathrm{G}_W(s)) &= \mathrm{e}^{a_1\left(\phi_1+\phi_2 s+(1-\phi_1-\phi_2)s^2-1\right)+a_2\left(\left(\phi_1+\phi_2 s+(1-\phi_1-\phi_2)s^2\right)^2-1\right)} \\
&= \mathrm{e}^{\left(a_1(\phi_1^2-1)+a_2(\phi_1^2-1)\right)+\left(a_1\phi_2+2\phi_1\phi_2\right)s+\left(a_1(1-\phi_1-\phi_2)+a_2(2\phi_1(1-\phi_1-\phi_2)+\phi_2^2)\right)s^2} \\
&\quad \mathrm{e}^{2a_2\phi_2(1-\phi_1-\phi_2)s^3+a_2(1-(\phi_1+\phi_2))^2 s^4},
\end{aligned}$$

which is the pgf of a 4th-Hermite with parameters $b_1 = a_1\phi_2 + 2\phi_1\phi_2$, $b_2 = a_1(1-\phi_1-\phi_2) + a_2(2\phi_1(1-\phi_1-\phi_2)+\phi_2^2)$, $b_3 = 2a_2\phi_2(1-\phi_1-\phi_2)$ and $b_4 = a_2(1-(\phi_1+\phi_2))^2$. Recalling that $\mathrm{G}_Z(s) = \mathrm{G}_X(s)/\mathrm{G}_X(\mathrm{G}_W(s))$, the marginal distribution of X_n cannot neither be a 2nd-Hermite since the parameter of s^4 is $-a_2(1-(\phi_1+\phi_2))^2 < 0$, but it should be positive to be the pgf of a 4th-Hermite distribution.

Other count distributions for X_n have considered such as the Binomial, Negative Binomial and Geometric. However, none of them leads to a proper known probability generating function and, hence, a known distribution for the innovations of the process in (15).

From another point of view, let $Z_n \sim \text{Poisson}(\lambda)$, then the distribution of X_n can be determined by:

$$\frac{G_X(s)}{G_X(\phi_1 + \phi_2 s + (1 - \phi_1 - \phi_2)s^2)} = \mathrm{e}^{\lambda(s-1)}. \tag{17}$$

However, any known probability generating function solves this equality. The latter means that, when innovations are Poisson, the marginal distribution of the model (15) is unknown.
In the same way, when $Z_n \sim \text{Hermite}(a_1, a_2)$, the marginal distribution of X_n is unknown since there is not a (known) probability generating function solving the following equality:

$$\frac{G_X(s)}{G_X(\phi_1 + \phi_2 s + (1 - \phi_1 - \phi_2)s^2)} = \mathrm{e}^{a_1(s-1)+a_2(s^2-1)}. \tag{18}$$

References

1. Fernández-Fontelo, A., Cabaña, A., Joe, H., Puig, P. and Moriña, D. Untangling serially dependent under-reported count data from gender-based violence. Under review.

2. Fernández-Fontelo, A., Cabaña, A., Puig, P. and Moriña, D. (2016). Under-reported data analysis with INAR-hidden Markov chains. Statistics in Medicine, 35: 4875-4890.

3. Kemp, C.D. and Kemp, A.W. (1965). Some properties of the "Hermite" distribution. Biometrika, 52: 381–394.