Series Approximations in Analysis of Risk

Costas A. Christophi[1] **Reza Modarres**[1]

The George Washington University, Department of Statistics
Washington, D.C. 20052, U.S.A.

Abstract. We consider a multiplicative risk function, $R = \prod_{i=1}^{p} x_i^{a_i}$, where x_i are positive random variables, independent but not identically distributed. We discuss and compare the simulated distribution of $S_p = \ln(R)$ with several asymptotic approximations. The Generalized Central Limit Theorem is used to obtain a normal approximation. The Edgeworth expansion of the distribution of S_p and the saddlepoint approximations are obtained. The accuracies of each of the above approximations are illustrated in several examples, where they are compared to the exact and the Monte Carlo results.

Keywords: Series Approximation; Edgeworth Expansion; Saddlepoint Method; Multiplicative Risk Model; Monte Carlo Simulation.

1. Introduction

A basic problem in risk assessment is to provide an accurate estimate and a measure of confidence in that estimate for different aspects of the risk function $R = h(\vec{X}) = h(x_1, \ldots, x_p)$. The joint occurrence of the risk factors is modeled by a p-dimensional probability distribution function $f_\theta(x_1, \ldots, x_p)$. The parameter θ is usually vector valued and the functional form of h is known. The estimation of the entire distribution function $G_R(r) = P(R \le r)$ and its inverse, the quantile function $Q_R(u) = G_R^{-1}(u)$ for $0 < u < 1$, are of primary concern and often, interest centers on accurate estimation of the upper or lower quantiles of R and their standard errors. In many cases of practical interest, risk is a function of several variables. It is also often the case that the data are only available separately for each of the risk factors; sometimes from different sources (surrogate or proxy data source, Hodges, 1987). A common strategy is to assume independence among the risk factors and consider the possible effects of dependence in a subsequent sensitivity analysis.

In many environmental applications the distribution of risk factors such as body weight, total skin area, concentration, inhalation, digestion, and

[1]This research was supported by a Cooperative Agreement with the Environmental Protection Agency, and a grant from the Water Research Unit of the University of District of Columbia. *E-mail*: coschri@gwu.edu, reza@gwu.edu.

consumption rates are positive and skewed to the right. A log transformation will then provide a more natural scale to analyze such measurements. In the multiplicative risk model the distribution of R is obtained from multiplication, division or exponentiation of the risk factors (Slob, 1994). If we let $R = \prod_{i=1}^{p} x_i^{a_i}$, then $R = x_1 \times x_2 \times \ldots \times x_p$ is the case considered in Hoffman and Hammonds (1994) with $a_i = 1$, $i = 1, \ldots, p$. Assuming $x_i > 0$, one can transform the variables to the log-scale and consider an additive risk model $S_p = \sum_{i=1}^{p} a_i \ln(x_i)$, which is generally more amenable to analysis.

Information on factors such as drinking water intake rate, soil ingestion rate, inhalation rate, dairy and fish intake rates is included in the Exposure Factors Handbook (USEPA, 1997) and is often utilized in multiplicative risk models to provide estimates of risk to human health as a result of exposure to chemicals or microbes. For example, USEPA (1992) computes Average Daily Dose (ADD) by averaging the Total Potential Dose (TPD) over Body Weight (BW) and an Averaging Time (AT); that is, $ADD = \frac{TPD}{BW \times AT}$. TPD is defined as $TPD = C \times IR \times ED$ where C is the contaminant concentration in the medium (air, food, soil, etc.), IR is the intake rate such as rate of inhalation, ingestion or dermal contact and ED is the exposure duration or the amount of time that the contaminated contact lasts; e.g. the length of stay in an area, the time spent showering, etc. We would like to characterize the distribution of ADD when distributional forms have been assumed for the individual risk factors such as BW, C and IR. As another example, consider characterizing the distribution of a Hazard Index (HI) for a specified chemical in consumed fish where $HI = \frac{C \times IR}{BW \times RfD}$. Here, C represents the concentration of a chemical contaminant in fish, IR is the ingestion rate of fish and RfD is the chemical-specific reference dose.

2. Methods of Solution

One can attempt to find an exact answer to the distribution of R, or approximate it using methods based on Monte Carlo simulation, or asymptotic approximations. The exact method uses the joint density function f and finds the distribution of R after a p-variate transformation followed by $p-1$ subsequent integrations. However, when $\vec{X}$ consists of several types of distributions, as it is usually the case in practice, the exact distribution of R is difficult to derive and Monte Carlo simulation is used to provide an approximation.

A Monte Carlo distribution of R can be obtained by generating a large number of replications of the random vector $\vec{X}$ from $f_\theta(\vec{x})$ and recording R for each. The resulting empirical distribution function for R serves as an estimate of its true distribution function. Ordered samples of R provide an estimator of the quantile function and the tail probabilities are estimated as the proportion of the $\vec{X}$ vectors that fall in the upper tail area. Monte Carlo simulation has been an effective tool in computing the distribution of R in most cases. It is possible to write routine programs to compute the Monte Carlo distribution of risk given several specified risk factors. A number of software packages (@Risk, 1996; Crystal Ball, 1997; Analytica, 2000; Decisionpro, 2000) have been developed to find the distribution of risk using Monte Carlo simulation.The main disadvantage of this technique is the number of replications necessary to obtain credible results. For example, computing $Q_R(u) = F_R^{-1}(u)$ for u close to zero or one may require a huge number of replications, each of which may require random number generation or resampling. Further discussion and illustrative examples, on this and subsequent sections, appear in Christophi and Modarres (2001).

3. Asymptotic Approximations

In this section we present the normal, Edgeworth and saddlepoint approximations to the distribution of

$$S_p = \ln(R) = \sum_{i=1}^{p} a_i \ln(x_i) = \sum_{i=1}^{p} y_i. \tag{1}$$

Asymptotic expansions allow us to expand the distribution function of S_p in terms of the normal distribution function and additional terms that depend on the moments of S_p and the normal density. We assume that the cumulative generating function of each risk factor is known. A standard assumption in Monte Carlo analysis of risk is knowledge of the form of the underlying distribution of risk. These assumptions are equivalent whenever the moments uniquely characterize the distribution.

One can justify the normal approximation of S_p based on the assertion that the risk factors are log-normal. Johnson et al. (1994) discuss the applications of the log-normal distribution to model positive random variables. Modarres, Nayak and Gastwirth (2002) assume that the risk factor y_i is such that the distribution of its logarithm belongs to a symmetric location-scale family and study the log-normal, log double exponential and log logistic distributions. If we assume y_i is normal, then the normal approximation of S_p will be exact. One can also note that S_p is the sum of p independent factors, which may not have identical distributions. The Central Limit Theorem, which justifies the use of this method of approximation, is then stated under different conditions. Generalized Central Limit Theorem (Serfling, 1980) states conditions under which this sum is approximately normal.

One may consider the normal approximation of S_p as the first term of a series and can often improve the normal approximation by considering higher order terms in the expansion of the characteristic function in the i.i.d. case. Heuristically, most series expansions seek to represent the distribution function of the variable of interest in terms of its moments and the distribution function of a target distribution, which is usually well-studied with known properties. When the target distribution is normal such an expansion uses the Hermite polynomials (see Barndorff-Nielsen and Cox, 1989).

Hall (1992) gives the following Edgeworth expansion for the average of p i.i.d. random variables $\{y_i\}$ with mean μ and variance σ^2

$$P\left(\frac{\sqrt{p}(\bar{Y}-\mu)}{\sigma} \leq w\right) = \Phi(w) + \phi(w)\left[\frac{-1}{6\sqrt{p}}\kappa(w^2-1) + O(p^{-1})\right] \quad (2)$$

where Φ and ϕ are the distribution and density function of a standard normal, respectively, and κ is the skewness, $E(Y-\mu)^3$. As noted (Goutis and Casella, 1999), the first term of the above formulation is the normal approximation, which is accurate to $O(p^{-1/2})$. When $\kappa = 0$ the accuracy improves to $O(p^{-1})$.

Let $y_1, y_2, \ldots, y_p$ be independent random variables with density f_{y_i}, mean μ_i, and variance σ_i^2. The moment generating function of y_i is $M_{y_i}(t) = E(\exp(ty_i))$ and its cumulative generating function is $K_{y_i}(t) = \ln(M_{y_i}(t))$. In order to obtain the Edgeworth expansion of $R = \prod_{i=1}^p x_i$ let $\rho_{r,i} = \kappa_{r,i}/\sigma_i^r$ denote the standardized cumulants of $\ln(x_i)$. We will use (**??**) to obtain an Edgeworth expansion to the distribution function of S_p. Let $\mu = E(S_p) = \sum_{i=1}^p \mu_i$, $\sigma^2 = Var(S_p) = \sum_{i=1}^p \sigma_i^2$, and $C_r = \sum_{i=1}^p \kappa_{r,i}/\sigma^r$. The formal inversion of $M_{S_p^*}(t)$ obtains $f_{S_p^*}(s_p^*)$ and by integration

$$F_{S_p^*}(s_p^*) \approx \Phi(s_p^*) - \phi(s_p^*)\left[\frac{1}{6}C_3H_2(s_p^*) + \frac{1}{24}C_4H_3(s_p^*) + \frac{1}{72}C_3^2H_5(s_p^*)\right] \quad (3)$$

where $S_p^* = \frac{S_p-\mu}{\sigma}$ and $H_r(x)$ is the Hermite polynomial of degree r. Note that $F_{S_p}(s_p) = F_{S_p^*}(s_p^*)$. The formal Edgeworth expansion requires that $\lambda_j = \kappa_{r,S_p}/\kappa_{2,S_p}^{3/2}$ for $j = 2, 3$ are bounded as $p \to \infty$.

One can improve the normal approximation to the distribution of the mean of a random sample by using the one-term or two-term Edgeworth expansion to correct for non-normality of the parent distribution. Consider (**??**) and change the variable to $y = \sigma w + \mu$. Differentiation yields

$$f_{\bar{y}}(y) = \frac{\sqrt{p}}{\sigma}\phi\Big(\frac{y-\mu}{\sigma/\sqrt{p}}\Big)\Bigg\{1 + \frac{\kappa}{6\sqrt{p}}\Big[\Big(\frac{y-\mu}{\sigma/\sqrt{p}}\Big)^3 - 3\Big(\frac{y-\mu}{\sigma/\sqrt{p}}\Big)\Big] + O(p^{-1})\Bigg\}. \quad (4)$$

Clearly, the normal approximation has error $O(p^{-1})$ when $\kappa = 0$. It also yields an order of accuracy of $O(p^{-1})$ for values of y near the mean as the term in bracket will be close to zero. One would expect better approximations near the center. This observation leads to the method of tilted Edgeworth expansion or saddlepoint approximation where a more accurate estimate for the distribution of a sum is obtained. Daniel (1954) describes an accurate approximation of the density of the mean of a sample through the direct inversion of the Fourier transform. Suppose $y_1, \ldots, y_p$ are i.i.d. with m.g.f. $M_y(t)$. The m.g.f. of the average $\bar{y}$ is $M_{\bar{y}}(t) = (M_y(t/p))^p$ and its c.g.f. is $K_{\bar{y}}(t) = pK_y(t/p)$. One can show that for large p

$$f_{\bar{y}}(\bar{y}) \approx (\frac{p}{2\pi K''(\hat{t}(\bar{y}))})^{1/2} \exp(p(K_y(\hat{t}(\bar{y})) - \hat{t}(\bar{y})\bar{y})). \quad (5)$$

where we choose $t = \hat{t}(\bar{y})$ such that it satisfies the saddlepoint equation $K'_{\bar{y}}(t) = \bar{y}$ and one can then integrate $f_{\bar{y}}$ to obtain an approximation to $F_{\bar{y}}$.

Lugannani and Rice (1980) provide a direct expansion of the integral as a series. Lugannani and Rice's formula uses the first term of the series $F_y(y) \approx \Phi(r) + \phi(r)[\frac{1}{r} - \frac{1}{q} + O(p^{-1})]$ where $r = sign[\hat{t}(y)][2(\hat{t}(y)y - K_y[\hat{t}(y)])]^{1/2}$ and $q = \hat{t}(y)(K''_y[\hat{t}(y)])^{1/2}$. Note that a more accurate approximation of $O(p^{-3/2})$ for the distribution at the origin exist (Jensen, 1995). Since $K_{S_p}(t) = \sum_{i=1}^{p} K_{y_i}(t)$ one can obtain a saddlepoint approximation for the distribution function of the sum of p independent but not identically distributed random variables. Solving $K'_{S_p}(t) = s_p$ for $t = \hat{t}(s_p)$ one obtains

$$F_{S_p}(s_p) \approx \Phi(r) + \phi(r)\left[\frac{1}{r} - \frac{1}{q}\right] \quad (6)$$

where $r = sign[\hat{t}(s_p)][2(\hat{t}(s_p)s_p - K_{S_p}[\hat{t}(s_p)])]^{1/2}$ and $q = \hat{t}(s_p)(K''_{S_p}[\hat{t}(s_p)])^{1/2}$.

4. Examples

Gamma distribution has been used in life testing, reliability theory and the theory of stochastic processes. This model allows us to consider the distribution of a variety of environmental factors that may enter the equation for risk. For example, consider the distribution of $S_p = \sum_{i=1}^{p} y_i$ where y_i is

$\Gamma(\alpha_i, \beta_i)$. Gamma distribution is a serious competing model any time log-normal or Weibull distributions are considered. The log-gamma distribution has also been used to model a variety of physical phenomenon.

Beta variables arise naturally as the distribution of an ordered variable from a uniform distribution. The beta family is multi-faceted and finds application in many fields. For interesting applications in risk analysis see Morgan and Henrion (1990). This distribution is often used to model proportions of contaminants in a medium; e.g. water since it is bounded by the same range as the distribution of concentrations. Ott (1995) discusses the modeling of dilution of pollutants in the environment using beta random variables and Johnson et al. (1994) discuss the use of log-beta in place of log-normal when modeling positively or negatively skewed data.

Example 1: This example considers approximating the distribution of the risk function $R = X$ where X has a log-gamma distribution with parameters α and β. From (1) we have $S_1 = \ln(R) = \ln(X) = Y$ has a gamma distribution. Let $\theta = E(Y) = \alpha\beta$ and $\sigma^2 = Var(Y) = \alpha\beta^2$. The normal approximation to the distribution function of Y is $\Phi(y^*)$ and its density is $\phi(y^*)$ where $y^* = (y - \theta)/\sigma$. The one and two-term Edgeworth approximations are obtained from

$$F_{Y^*}(y^*) \approx \Phi(y^*) - \phi(y^*)\left[\frac{1}{6}\kappa_{3,1}H_2(y^*) + \frac{1}{24}\kappa_{4,1}H_3(y^*) + \frac{1}{72}\kappa_{3,1}^2 H_5(y^*)\right]$$

where $\kappa_{r,S_p} = \sum_{i=1}^{p} \kappa_{r,y_i}$ and $\kappa_{r,1} = \alpha\Gamma(r)\beta^r$. The saddlepoint solution is obtained by solving the equation $K_y'(t) = \frac{\alpha\beta}{1-\beta t} = y$ for $t < 1/\beta$. The saddlepoint solution is $\hat{t}(y) = 1/\beta - \alpha/y$. Thus, we obtain $F_Y(y) \approx \Phi(r) + \phi(r)[\frac{1}{r} - \frac{1}{q}]$ as an approximation to the c.d.f. of Y where $r = sign(1/\beta - \alpha/y)[2(y/\beta - \alpha + \alpha\ln(\frac{\alpha\beta}{y}))]^{1/2}$ and $q = (y/\beta - \alpha)/\sqrt{\alpha}$.

Figures 1a-1c show how well the approximations perform for three distinct distributional shapes, $\alpha = 0.5, 1.0$ and 1.5, respectively. Since β is a scaling factor it does not affect the relative comparisons and is set to one. Since $p = 1$ is small, the supports of the exact and approximating distributions are not the same and we observe poor performance in the lower tail. Figure 1a indicates that the normal and Edgeworth expansions should not be used in this case. All approximations converge to the exact values as α becomes larger. This effect is observed by comparing Figures 1a to 1c. Note that the saddlepoint approximation is indistinguishable from the exact result in these cases. Using Stirling's approximation, it can be shown that the saddlepoint approximation is exact in this case.

Example 2: This example considers the convolution of $y_1 = \ln x_1$ and $y_2 = \ln x_2$ where x_1 has a normal distribution $N(\mu, \sigma^2)$ and x_2 has a gamma distribution $\Gamma(\alpha, \beta)$. One can show $K_{S_2}(t) = -\alpha \ln(1-\beta t) + \mu t + \sigma^2 t^2/2$. The saddlepoint solution is obtained by solving the saddlepoint equation $K'_{S_2}(t) = \frac{\alpha\beta}{1-\beta t} + \mu + \sigma^2 t = s_2$ for $t < 1/\beta$. The saddlepoint solution $\hat{t}(s_2)$ is obtained as the smaller root of the quadratic equation $\beta\sigma^2 t^2 - (\sigma^2 - \beta\mu + \beta s_2)t - (\alpha\beta + \mu - s_2) = 0$. The mean and variance of S_2 are $\theta = \alpha\beta + \mu$ and $\delta^2 = \alpha\beta^2 + \sigma^2$, respectively. Its r-*th* cumulant for $r > 2$ is $\Gamma(r)\alpha\beta^r$. The exact distribution of S_2 is based on a million simulated values. We take the percentage points of the simulated distribution as the percentile at which approximations are evaluated. The normal approximation of the distribution function of S_2 is $\Phi(s_2{}^*)$ where $s_2^* = (s_2 - \theta/)\delta$. The Edgeworth approximation is given by (**??**). For the saddlepoint approximation to the c.d.f. of s_2, we use $F_{s_2}(s_2) \approx \Phi(r) + \phi(r)[\frac{1}{r} - \frac{1}{q}]$ where r and q are obtained from (**??**). Figure 2a compares the distributions when $(\alpha, \beta) = (1, 1)$ and $(\mu, \sigma^2) = (1, 0.1)$. Figure 2b displays the result for $(\alpha, \beta) = (0.5, 10)$ and $(\mu, \sigma^2) = (0, 1)$. Notice that the two-term Edgeworth approximation is negative at the extreme loer tail.

Example 3: This example considers the distribution of the risk function

$$R = \frac{x_3 x_5}{x_1 x_2 x_4}.$$

This function has the same form as ADD, which was defined in the introduction. Let $S_p = -\ln(x_1) - \ln(x_2) + \ln(x_3) - \ln(x) + \ln(x_5) = \sum_{i=1}^{5} y_i$. Here, we assume that $y_1 = -\ln(x_1)$ has an exponential distribution with mean $\eta = 1$, $y_2 = -\ln(x_2) \sim \beta(\alpha_2, \beta_2)$, and $y_3 = \ln(x_3) \sim \text{Normal}(\mu, \sigma^2)$. We also assume that $y_4 = \ln(x_4)$ has a triangular distribution with minimum a, mode b and maximum c, and that $y_5 = \ln(x_5) \sim \Gamma(\alpha_5, \beta_5)$. The rth cumulant of y_1 is $\kappa_{r,1} = (r-1)!$, and for y_2 is $\kappa_{r,2} = (r-1)! \sum_{j=0}^{\beta_2 - 1} (\alpha_2 + j)^{-r}$ if β_2 is an integer. We will fix the scale parameter β_2 of x_2 to one and consider a variety of shapes for x_2 through α_2. To simplify we will consider two special cases for x_4. Case I assumes $a = 0$ and $b = c = 1$; i.e. $y_4 \sim \beta(\alpha_4 = 2, \beta_4 = 1)$ with $\kappa_{r,4} = (r-1)!/2^r$ and case II assumes $a = b = 0$ and $c = 1$; $y_4 \sim \beta(\alpha_4 = 1, \beta_4 = 2)$ with $\kappa_{r,4} = (r-1)!(1 + 1/2^r)$. The first two cumulants of y_3 are μ and σ^2, respectively. Its remaining cumulants are zero. For x_5 we have $\kappa_{r,5} = (r-1)!\alpha_5\beta_5^r$.

The c.g.f.'s of $y_1, \ldots y_5$ are $K_{y_1}(t) = -\ln(1-t)$ for $t < 1$, $K_{y_2}(t) = \log \alpha_2/(\alpha_2 - t)$ for $t < \alpha_2$, $K_{y_3}(t) = \mu t + \frac{1}{2}\sigma^2 t^2$, $K_{y_4}(t) = \log 2/(2-t)$ for $t < 2$ when $a = 0$ and $b = c = 1$, $K_{y_4}(t) = \log 2/(1-t)(2-t)$ for $t < 1$ when

$a = b = 0$ and $c = 1$, and $K_{y_5}(t) = -\alpha^* \log(1-\beta^* t)$ for $t < 1\beta^*$, respectively. Since $\kappa_{r,S_p} = \sum_{i=1}^{5} \kappa_{r,y_i}$ the mean and variance of S_p are determined from $\mu_{S_p} = \kappa_{1,S_p}$ and $\sigma^2_{S_p} = \kappa_{2,S_p}$, respectively. The exact distribution of S_p is based on a million simulated values. The normal approximation of the distribution function of S_p is $\Phi({s_p}^*)$ and its density is $\phi(s^*_p)$ where $s^*_p = (S_p - \mu_{S_p})/\sigma_{S_p}$. The Edgeworth approximation is given by (**??**) where C's are determined by the cumulants. The saddlepoint solution is obtained by solving the saddlepoint equation $K'_{S_p}(t) = s_p$ for $t < min(1, \alpha_2, 1/\beta^*)$. The saddlepoint solution $\hat{t}(s_p)$ is obtained as the smallest real root of a fourth-degree polynomial in each case. Thus, for an approximation to the c.d.f. of S_p we use $F_{S_p}(s_p) \approx \Phi(r) + \phi(r)[\frac{1}{r} - \frac{1}{q}]$ where r and q are obtained from (**??**). The exact distribution of S_5 is based on a million simulated values. We take the percentage points of the simulated distribution as the percentile at which approximations are evaluated. Figure 3a compares the distribution of S_5 under case I where $\eta = 1,\ \alpha_2 = .1,\ \beta_2 = 1,\ \mu = 0,\ \sigma^2 = 1,\ a = 0,\ b = 1,\ c = 1,\ \alpha^* = 1$ and $\beta^* = 1$. Figure 3b compares the distribution of S_5 under case II where $\eta = 1,\ \alpha_2 = 1,\ \beta_2 = 1,\ \mu = 0,\ \sigma^2 = 1,\ a = 0,\ b = 0,\ c = 1,\ \alpha^* = 1$, and $\beta^* = 1$.

5. Discussion

The National Academy of Sciences (NRC, 1994) and USEPA (1997) have stressed the need to distinguish between variability and uncertainty in the analysis of risk. A two-dimensional Monte Carlo strategy has been advocated (Hoffman and Hammonds, 1994; Burmaster and Wilson, 1996) and mixture distributions have been used to justify this approach. Note that Nayak and Kundu (2001) argue that the analysis is only relevant if the results are interpreted in a Bayesian context. The process can be extremely time consuming and to speed up computations it is desired to replace the computer intensive part of the process by an asymptotic approxiamtion.

It is noted that when the supports of the exact and approximating distributions are not the same there is poor performance in the lower tail of normal and Edgeworth approximations. Normal approximation underestimates the risk in the central region of the distribution within the interquartile range. Edgeworth approximations perform better than normal in the interquartile range. However, they underestimate or overestimate R in the upper tail.

Moreover, Edgeworth expansion provides an absolute error. Since interest centers on accurate estimation of the upper or lower quantiles of the risk function we look for approximations with a small relative error. Saddlepoint approximation will always produce a positive approximation over the range of the sample mean.

Acknowledgment. We would like to express our appreciation to Professors Hosam M. Mahmoud and Tapan K. Nayak for stimulating discussions.

References

[1] Analytica, (2000). Lumina Decision Systems, Los Gatos, CA, (http://www.lumina.com).

[2] Barndorff–Nielsen, O.E. and Cox, D.R. (1989). Asymptotic Techniques for Use in Statistics. Chapman and Hall.

[3] Burmaster, D.E. and Wilson, A. M. (1996). An Introduction to Second-Order Random Variables in Human Health Risk Assessment. Human and Ecological Risk Assessment, 2(4), 892-919.

[4] Christophi, C. A. and Modarres, R. (2001). Series Approximations in Analysis of Risk. Technical Report. Department of Statistics. The George Washington University.

[5] Crystal Ball (1997). Decisioneering, Inc. Denver, Colorado, (http://www.decisioneering.com).

[6] Daniels, H.E. (1954). Saddlepoint Approximations in Statistics. Annals of Mathematical Statistics, 25, 631–650.

[7] Decisionpro (2000). Vanguard Software Corp, Cary, NC, (http://www.vanguardsw.com).

[8] Goutis, C. and Casella, G. (1999). Explaining the Saddlepoint Approximation. The American Statistician, V. 53, No, 3, 216–224.

[9] Hall, P. (1992). The Bootstrap and Edgeworth Expansion. Springer-Verlag.

[10] Hodges, James S. (1987). Uncertainty, Policy Analysis and Statistics, *Statistical Science*, Vol. 2, No. 3, 259-275.

[11] Hoffman, F.O. and Hammonds, J.S. (1994). Propagation of Uncertainty in Risk Assessment: The Need to distinguish between Uncertainty due to Lack of Knowledge and Uncertainty due to Variability. Risk Analysis, 14, 707-712.

[12] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). Continuous Univariate Distributions, 2nd edition. John Wiley & Sons, Inc.

[13] Jensen, J. L. (1995). Saddlepoint Approximations. Oxford Press.

[14] Lugannani, R. and Rice, S. O. (1980). Saddlepoint Approximation for the Distribution of the Sum of Independent random variables. Adv. Appl. Prob., 12, 475-490.

[15] Modarres, R, Nayak, T. N. and Gastwrith, J. L. (2002). Estimation of Upper Quantiles of Risk Factors Under Model and Parameter Uncertainty. To appear in Computational Statistics and Data Analysis.

[16] Morgan, M. G. and Henrion, M. (1990). Uncertainty: A guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambrige University Press.

[17] National Research Council (NRC). (1994). Science and Judgment in Risk Assessment. National Academy Press, Washington DC.

[18] Nayak, T. K. and Kundu, S. (2001). Calculating and Describing Uncertainty in Risk Assessment: The Bayesian Approach. Journal of Human and Ecological Risk Assessment, Vol. 7, no. 2, 307-328.

[19] Ott, Wayne R. (1995). Environmental Statistics and Data Analysis, Lewis Publishers, New York.

[20] @Risk Software (1996). User's Guide, Palisade Corporation. Newfield, New York (http://www.palisade.com).

[21] Serfling, R. J. (1980). Approximation Theorems of Mathematical Statistics. John Wiley & Sons.

[22] Slob, W. (1994). Uncertainty Analysis in Multiplicative Models. Risk Analysis, No. 4, 571–576.

[23] USEPA (1992). Guidelines for Exposure Assessment. Washington DC: Office of Research and Development. EPA/600/2-92/001.

[24] USEPA (1997). Exposure Factors Handbooks. Volumes I-III. Washington DC: Office of Research and Development. EPA/600/P-95/00Fa, EPA/600/P-95/00Fb, and EPA/600/P-95/00Fc.