9

Generalized Poisson Distribution

9.1 Introduction and Definition

Let X be a discrete r.v. defined over nonnegative integers and let $P_x(\theta, \lambda)$ denote the probability that the r.v. X takes a value x. The r.v. X is said to have a GPD with parameters θ and λ if

$$P_x(\theta, \lambda) = \begin{cases} \theta(\theta + \lambda x)^{x-1} e^{-\theta-\lambda x}/x!, & x = 0, 1, 2, \ldots, \\ 0, & \text{for } x > m \quad \text{if} \quad \lambda < 0, \end{cases} \tag{9.1}$$

and zero otherwise, where $\theta > 0$, $\max(-1, \ -\theta/m) \leq \lambda \leq 1$, and m (≥ 4) is the largest positive integer for which $\theta + m\lambda > 0$ when $\lambda < 0$. The parameters θ and λ are independent, but the lower limits on λ and $m \geq 4$ are imposed to ensure that there are at least five classes with nonzero probability when λ is negative. The GPD model reduces to the Poisson probability model when $\lambda = 0$. Consul and Jain (1973a, 1973b) defined, studied, and discussed some applications of the GPD in (9.1).

The GPD belongs to the class of Lagrangian distributions $L(f; g; x)$, where $f(z) = e^{\theta(z-1)}$, $\theta > 0$, and $g(z) = e^{\lambda(z-1)}$, $0 < \lambda < 1$, and is listed as (6) in Table 2.3. It belongs to the subclass of MPSD. Naturally, it possesses all the properties of these two classes of distributions.

When λ is negative, the model includes a truncation due to $P_x(\theta, \lambda) = 0$ for all $x > m$ and the sum $\sum_{x=0}^{m} P_x(\theta, \lambda)$ is usually a little less than unity. However, this truncation error is less than 0.5% when $m \geq 4$ and so the truncation error does not make any difference in practical applications.

The multiplication of each $P_x(\theta, \lambda)$ by $[F_m(\theta, \lambda)]^{-1}$, where

$$F_m(\theta, \lambda) = \sum_{x=0}^{m} P_x(\theta, \lambda), \tag{9.2}$$

has been suggested for the elimination of this truncation error. (See Consul and Shoukri (1985), Consul and Famoye (1989b).)

Lerner, Lone, and Rao (1997) used analytic functions to prove that the GPD will sum to 1. Tuenter (2000) gave a shorter proof based upon an application of Euler's difference lemma.

The properties and applications of the GPD have been discussed in full detail in the book *Generalized Poisson Distribution: Properties and Applications*, by Consul (1989a). Accordingly, some important results only are being given in this chapter. The GPD and some of its properties have also been described in *Univariate Discrete Distributions* by Johnson, Kotz, and Kemp (1992).

9.2 Generating Functions

The pgf of the GPD with parameters (θ, λ) is given by the Lagrange expansion in (1.78) as

$$G(u) = e^{\theta(z-1)}, \text{ where } z = u\, e^{\lambda(z-1)}. \tag{9.3}$$

The above pgf can also be stated in the form

$$G(u) = e^{\theta(w(u)-1)}, \tag{9.4}$$

where the function $w(u)$ is defined by the relation

$$w(u) = u \ \exp\{\lambda(w(u) - 1)\}. \tag{9.5}$$

The function $w(u)$ is 0 at $u = 0$ and 1 at $u = 1$, and its derivative is

$$w'(u) = \left[e^{-\lambda(w(u)-1)} - u\lambda\right]^{-1}. \tag{9.6}$$

By putting $z = e^s$ and $u = e^\beta$ in (9.3), one obtains the mgf for the GPD model as

$$M_x(\beta) = e^{\theta(e^s-1)}, \text{ where } s = \beta + \lambda(e^s - 1). \tag{9.7}$$

Thus, the cgf of the GPD becomes

$$\psi(\beta) = \ln M_x(\beta) = \theta(e^s - 1), \text{ where } s = \beta + \lambda(e^s - 1). \tag{9.8}$$

It has been shown in chapter 2 that the GPD is a particular family of the class of Lagrangian distributions $L(f; g; x)$ and that the mean μ and variance σ^2 are

$$\mu = \theta(1 - \lambda)^{-1}, \qquad \sigma^2 = \theta(1 - \lambda)^{-3}. \tag{9.9}$$

The variance σ^2 of the GPD is greater than, equal to, or less than the mean μ according to whether $\lambda > 0$, $\lambda = 0$, or $\lambda < 0$, respectively.

Ambagaspitiya and Balakrishnan (1994) have expressed the pgf $P_x(z)$ and the mgf $M_x(z)$ of the GPD (9.1) in terms of Lambert's W function as

$$M_x(z) = \exp\left\{-(\lambda/\theta)\left[W(-\theta\exp(-\theta - z)) + \theta\right]\right\}$$

and

$$P_x(z) = \exp\left\{-(\lambda/\theta)\left[W(-\theta z\exp(-\theta)) + \theta\right]\right\},$$

where W is the Lambert's function defined as

$$W(x)\exp(W(x)) = x.$$

They have derived the first four moments from them, which are the same as those given in (9.9) and (9.13).

9.3 Moments, Cumulants, and Recurrence Relations

All the cumulants and moments of the GPD exist for $\lambda < 1$. Consul and Shenton (1975) and Consul (1989a) have given the following recurrence relations between the noncentral moments μ'_k and the cumulants K_k:

$$(1-\lambda)\ \mu'_{k+1} = \theta\ \mu'_k + \theta\frac{\partial \mu'_k}{\partial\theta} + \lambda\frac{\partial \mu'_k}{\partial\lambda}, \qquad k = 0, 1, 2, \ldots, \tag{9.10}$$

$$(1-\lambda)K_{k+1} = \lambda\frac{\partial K_k}{\partial\lambda} + \theta\frac{\partial K_k}{\partial\theta}, \qquad k = 1, 2, 3, \ldots. \tag{9.11}$$

A recurrence relation between the central moments of the GPD is

$$\mu_{k+1} = \frac{k\theta}{(1-\lambda)^3}\ \mu_{k-1} + \frac{1}{1-\lambda}\left\{\frac{d\ \mu_k(t)}{dt}\right\}_{t=1}, \qquad k = 1, 2, 3, \ldots, \tag{9.12}$$

where $\mu_k(t)$ is the central moment μ_k with θ and λ replaced by θt and λt, respectively. The mean and variance of GPD are given in (9.9). Some other central moments of the model are

$$\left.\begin{aligned}
&\mu_3 = \theta(1+2\lambda)(1-\lambda)^{-5},\\
&\mu_4 = 3\theta^2(1-\lambda)^{-6} + \theta(1+8\lambda+6\lambda^2)(1-\lambda)^{-7},\\
&\mu_5 = 10\theta^2(1+2\lambda)(1-\lambda)^{-8} + \theta(1+22\lambda+58\lambda^2+24\lambda^3)(1-\lambda)^{-9},\\
&\text{and}\\
&\mu_6 = 15\theta^3(1-\lambda)^{-9} + 5\theta^2(5+32\lambda+26\lambda^2)(1-\lambda)^{-10}\\
&\qquad + \theta(1+52\lambda+328\lambda^2+444\lambda^3+120\lambda^4)(1-\lambda)^{-11}.
\end{aligned}\right\} \tag{9.13}$$

By using the values of $\mu_2 = \sigma^2$ in (9.9), μ_3 and μ_4 in (9.13), the expressions for the coefficients of skewness (β_1) and kurtosis (β_2) are given by

$$\beta_1 = \frac{1+2\lambda}{\sqrt{\theta(1-\lambda)}} \quad \text{and} \quad \beta_2 = 3 + \frac{1+8\lambda+6\lambda^2}{\theta(1-\lambda)}. \tag{9.14}$$

For any given value of λ, the skewness of the GPD decreases as the value of θ increases and becomes zero when θ is infinitely large. Also, for any given value of θ, the skewness is infinitely large when λ is close to unity. The skewness is negative for $\lambda < -\frac{1}{2}$.

For all values of θ and for all values of λ in $0 < \lambda < 1$, the GPD is leptokurtic as $\beta_2 > 3$. When

$$-\frac{1}{6}\sqrt{10} - \frac{2}{3} < \lambda < \frac{1}{6}\sqrt{10} - \frac{2}{3},$$

the GPD becomes platykurtic since β_2 becomes less than 3.

The expressions for the mean deviation, the negative integer moments, and the incomplete moments are all given in the book by Consul (1989a).

9.4 Physical Models Leading to GPD

The GPD does relate to a number of scientific problems and can therefore be used to describe many real world phenomena.

Limit of Generalized Negative Binomial Distribution

The discrete probability model of GNBD, discussed in chapter 10, is given by

$$P(X=x)=\frac{m}{m+\beta x}\binom{m+\beta x}{x}p^x(1-p)^{m+\beta x-x}, \qquad x=0,1,2,\ldots, \tag{9.15}$$

and zero otherwise, where $0<p<1$ and $1\le\beta<p^{-1}$ for $m>0$.

Taking m and β to be large and p to be small such that $mp=\theta$ and $\beta p=\lambda$, the GNBD approaches the GPD.

Limit of Quasi-Binomial Distribution

While developing urn models dependent upon predetermined strategy, Consul (1974) obtained a three-parameter QBD-I defined in (4.1). If $p\to 0,\ \phi\to 0$, and n increases without limit such that $n\,p=\theta$ and $n\,\phi=\lambda$, the QBD-I approaches the GPD.

Consul and Mittal (1975) gave another urn model which provided a three-parameter QBD-II defined in (4.73). When $p\to 0,\ \alpha\to 0$ and n increases without limit such that $n\,p=\theta$ and $n\alpha=\lambda$, the QBD-II approaches the GPD.

Limit of Generalized Markov–Pólya Distribution

Janardan (1978) considered a four-urn model with predetermined strategy and obtained the generalized Markov–Pólya distribution given by

$$P(X=k)=\frac{pq(1+Ns)\binom{N}{k}\prod_{j=0}^{k-1}(\theta+ks+jr)\prod_{j=0}^{N-k-1}(q+Ns-ks+jr)}{(p+ks)(q+Ns-ks)\prod_{j=0}^{N-1}(q+Ns+jr)} \tag{9.16}$$

for $k=0,1,2,\ldots,N$ and zero otherwise, where $0<p<1,\ q=1-p,\ r>0,\ s>0$, and N is a positive integer.

When N increases without limit and $p\to 0,\ r\to 0,\ s\to 0$ such that $Np=\theta,\ Ns=\lambda$, and $Nr\to 0$, the Markov–Pólya distribution in (9.16) approaches the GPD.

Models Based on Difference-Differential Equations

Consul (1988) provided two models, based on difference-differential equations, that generate the GPD model. Let there be an infinite but countable number of available spaces for bacteria or viruses or other micro-organisms. Let the probability of finding x micro-organisms in a given space be $P_x(\theta,\lambda)$.

Model I. Suppose the mean $\mu(\theta,\lambda)$ for the probability distribution of the micro-organisms is increased by changing the parameter θ to $\theta+\Delta\theta$ in such a way that

$$\frac{dP_0(\theta,\lambda)}{d\theta}=-P_0(\theta,\lambda), \tag{9.17}$$

and

$$\frac{dP_x(\theta,\lambda)}{d\theta}=-P_x(\theta,\lambda)+P_{x-1}(\theta+\lambda,\lambda), \tag{9.18}$$

for all integral values of $x>0$ with the initial conditions $P_0(0,\lambda)=1$ and $P_x(0,\lambda)=0$ for $x>0$, then the probability model $P_x(\theta,\lambda)$, $x=0,1,2,\ldots$, is the GPD.

Proof. See Consul (1989a).

Model II. Suppose the mean $\mu(\theta, \lambda)$ for the distribution of the micro-organisms is increased by changing the parameter λ to $\lambda + \Delta\lambda$ in such a way that

$$\frac{dP_0(\theta, \lambda)}{d\lambda} = 0 \tag{9.19}$$

and

$$\frac{dP_x(\theta, \lambda)}{d\lambda} = -xP_x(\theta, \lambda) + \frac{(x-1)\theta}{\theta + \lambda} P_{x-1}(\theta + \lambda, \lambda) \tag{9.20}$$

for all integral values of $x > 0$ with the initial conditions $P_x(\theta, 0) = e^{-\theta}\,\theta^x/x!$ for all values of x. Then the probability given by $P_x(\theta, \lambda), x = 0, 1, 2, \dots$ is the GPD.

Proof. See Consul (1989a).

Queuing Process

Let $g(z)$, the pgf of a Poisson distribution with mean λ, denote the pgf of the number of customers arriving for some kind of service at a counter and let X be a random variable which denotes the number of customers already waiting for service at the counter before the service begins. Also, let $f(z)$, the pgf of another Poisson distribution with mean θ, denote the pgf of X. Consul and Shenton (1973a) showed that the number of customers Y served in a busy period of the counter is a GPD. (See more on queuing process in chapter 6.)

Branching Process

Suppose

(a) the total number of units in a group is large,
(b) the probability of acquiring a particular characteristic by a unit in the group is small,
(c) each of the units having the particular characteristic becomes a spreader of the characteristic for a short time, and
(d) the number of members in the group where each spreader having the particular characteristic is likely to spread it is also large.

Consul and Shoukri (1988) proved that the total number of individuals having the particular characteristic (i.e., the total progeny in the branching process) is the GPD. (See more on the branching process in chapter 6.)

Thermodynamic Process

Consul (1989a) described generating GPD from a thermodynamic process with forward and backward rates. Let the forward and backward rates be given, respectively, by

$$a_k = (\theta + k\lambda)^{1-k} \quad \text{and} \quad b_k = ke^{\lambda}(\theta + k\lambda)^{1-k}, \tag{9.21}$$

which become smaller and smaller in value as k increases. Under the above forward and backward rates, the steady state probability distribution of a first-order kinetic energy process is that of the GPD model.

Proof. See Consul (1989a).

9.5 Other Interesting Properties

Theorem 9.1 (Convolution Property). *The sum $X + Y$ of two independent GP variates X and Y, with parameters (θ_1, λ) and (θ_2, λ), respectively, is a GP variate with parameters $(\theta_1 + \theta_2, \lambda)$ (Consul, 1989a).*

Theorem 9.2 (Unimodality). *The GPD models are unimodal for all values of θ and λ and the mode is at $x = 0$ if $\theta e^{-\lambda} < 1$ and at the dual points $x = 0$ and $x = 1$ when $\theta e^{-\lambda} = 1$, and for $\theta e^{-\lambda} > 1$ the mode is at some point $x = M$ such that*

$$\left(\theta - e^{-\lambda}\right)\left(e^{\lambda} - 2\lambda\right)^{-1} < M < a, \tag{9.22}$$

where a is the smallest value of M satisfying the inequality

$$\lambda^2 M^2 + M\left[2\lambda\theta - (\theta + 2\lambda)\, e^{\lambda}\right] + \theta^2 > 0 \tag{9.23}$$

(Consul and Famoye (1986a)).

A number of useful relations on the derivatives, integrals, and partial sums of the GPD probabilities are given in the book by Consul (1989a).

9.6 Estimation

Let a random sample of n items be taken from the GPD and let $x_1, x_2, \ldots, x_n$ be their corresponding values. If the sample values are classified into class frequencies and n_i denotes the frequency of the ith class, the sample sum y can be written as

$$y = \sum_{j=1}^{n} x_j = \sum_{i=0}^{k} i n_i, \tag{9.24}$$

where k is the largest of the observations, $\sum n_i = n$, and $\bar{x} = y/n$ is the sample mean. The sample variance is given by

$$s^2 = (n-1)^{-1} \sum_{i=0}^{k} n_i (i - \bar{x})^2 = (n-1)^{-1} \sum_{j=1}^{n} (x_j - \bar{x})^2. \tag{9.25}$$

9.6.1 Point Estimation

Moment Estimation

Consul and Jain (1973a) gave the moment estimators in the form

$$\tilde{\theta} = \sqrt{\frac{\bar{x}^3}{s^2}} \qquad \text{and} \qquad \tilde{\lambda} = 1 - \sqrt{\frac{\bar{x}}{s^2}}\,. \tag{9.26}$$

Shoukri (1980) computed the asymptotic biases and the asymptotic variances of the moment estimators correct up to the second order of approximation. They are

$$b(\tilde{\theta}) \simeq \frac{1}{4n}\left[5\theta + \frac{3\lambda\,(2+3\lambda)}{1-\lambda}\right], \tag{9.27}$$

$$b(\tilde{\lambda}) \simeq -\frac{1}{4n\theta}\left[5\theta(1-\lambda) + \lambda\left(10+9\lambda^2\right)\right], \tag{9.28}$$

$$V(\tilde{\theta}) \simeq \frac{\theta}{2n}\left[\theta + \frac{2-2\lambda+3\lambda^2}{1-\lambda}\right], \tag{9.29}$$

$$V(\tilde{\lambda}) \simeq \frac{1-\lambda}{2n\theta}\left[\theta - \theta\lambda + 2\lambda + 3\theta^2\right], \tag{9.30}$$

and

$$Cov(\tilde{\theta},\ \tilde{\lambda}) \simeq -\frac{1}{2n}\left[\theta(1-\lambda) + 3\lambda^2\right]. \tag{9.31}$$

Bowman and Shenton (1985) stated that the ratio of the sample variance to the sample mean estimates a simple function of the GPD dispersion parameter λ. They provided moment series to order n^{-24} for related estimators and obtained exact integral formulations for the first two moments of the estimator.

Estimation Based on Sample Mean and First (Zero-Class) Frequency

When the frequency for the zero-class in the sample is larger than most other class frequencies or when the graph of the sample distribution is approximately L-shaped (or reversed J-shaped), estimates based upon the mean and zero-class frequency may be appropriate. These estimates are given by

$$\theta^* = -\ln(n_0/n) \qquad \text{and} \qquad \lambda^* = 1 - \theta^*/\bar{x}. \tag{9.32}$$

Their variances and covariance up to the first order of approximation are

$$V(\theta^*) \simeq \frac{1}{n}\left(e^\theta - 1\right), \tag{9.33}$$

$$V(\lambda^*) \simeq \frac{1-\lambda}{n\theta^2}\left[(1-\lambda)\left(e^\theta - 1\right) + \theta(2\lambda-1)\right], \tag{9.34}$$

and

$$Cov(\theta^*, \lambda^*) \simeq -\frac{1-\lambda}{n\theta}\left(e^\theta - \theta - 1\right). \tag{9.35}$$

Maximum Likelihood Estimation

The ML estimate $\hat{\lambda}$ of λ is obtained by solving

$$H(\lambda) = \sum_{x=0}^{k} \frac{x(x-1)n_x}{\bar{x} + (x-\bar{x})\lambda} - n\bar{x}. \tag{9.36}$$

The ML estimate $\hat{\theta}$ of θ is obtained from

$$\hat{\theta} = \bar{x}(1-\hat{\lambda}).$$

Consul and Shoukri (1984) proved that the ML estimates $\hat{\theta} > 0$ and $\hat{\lambda} > 0$ are unique when the sample variance is larger than the sample mean. Consul and Famoye (1988) showed that if the sample variance is less than the sample mean, the ML estimates $\hat{\theta} > 0$ and $\hat{\lambda} < 0$ are also unique.

The GPD satisfies the regularity conditions given by Shenton and Bowman (1977). The variances and covariance of the ML estimators up to the first order of approximation are

$$V(\hat{\theta}) \simeq \frac{\theta(\theta+2)}{2n}, \tag{9.37}$$

$$V(\hat{\lambda}) \simeq \frac{(\theta+2\lambda-\theta\lambda)(1-\lambda)}{2n\theta}, \tag{9.38}$$

and

$$Cov(\hat{\theta}, \hat{\lambda}) \simeq -\frac{\theta(1-\lambda)}{2n}. \tag{9.39}$$

Consul and Shoukri (1984) gave the asymptotic biases as

$$b(\hat{\theta}) \simeq \frac{-\theta(5\theta^2+28\theta\lambda-6\theta\lambda^2+24\lambda^2)}{2n(1-\lambda)(\theta+2\lambda)^2(\theta+3\lambda)} \tag{9.40}$$

and

$$b(\hat{\lambda}) \simeq \frac{5\theta^3(1-\lambda)-2\theta^2\lambda\left(2\lambda^2+9\lambda-13\right)+4\theta\lambda^2(11-6\lambda)+24\lambda^2}{2n(1-\lambda)\theta(\theta+2\lambda)^2(\theta+3\lambda)}. \tag{9.41}$$

In comparing other estimators with the ML estimators, Consul (1989a) gave the asymptotic relative efficiency (ARE). The ARE for the moment estimators $\tilde{\theta}$ and $\tilde{\lambda}$ is given by

$$\text{ARE}(\tilde{\theta}, \tilde{\lambda}) = 1 - \frac{3\lambda^2}{\theta(1-\lambda)+\lambda(2+\lambda)}. \tag{9.42}$$

The ARE in (9.42) decreases monotonically as the value of λ increases, while it increases monotonically with θ. It was suggested that the moment estimators were reliable when $-0.5 < \lambda < 0.5$ and $\theta > 2$.

The ARE for the estimators based on mean and zero-class frequency is given by

$$\text{ARE}(\theta^*, \lambda^*) = \frac{\lambda+\theta/2}{\lambda+\left(e^{\theta}-1-\theta\right)\theta^{-1}} < 1. \tag{9.43}$$

For small values of θ, the estimators based on the mean and zero-class frequency will be better than the moment estimators.

Empirical Weighted Rates of Change Estimation

Famoye and Lee (1992) obtained point estimates for parameters θ and λ by using the empirical weighted rates of change (EWRC) method. Let $f_x = n_x/n$ be the observed frequency proportion for class x. The GPD likelihood equations can be written as

$$\sum_x f_x \frac{\partial}{\partial\theta_i} \ln P_x(\theta, \lambda) = 0, \quad i = 1, 2,$$

where $\theta_1 = \theta$ and $\theta_2 = \lambda$. From the fact that $\sum_x P_x(\theta, \lambda) = 1$, we obtain

$$\sum_x P_x \frac{\partial}{\partial \theta_i} \ln P_x(\theta, \lambda) = 0.$$

On combining the above with the likelihood equations, Famoye and Lee (1992) obtained the weighted discrepancies estimating equations as

$$\sum_x [f_x - P_x(\theta, \lambda)] \left[\frac{x(\theta + \lambda)}{\theta(\theta + \lambda x)} - 1 \right] = 0$$

and

$$\sum_x [f_x - P_x(\theta, \lambda)] \left[\frac{x(x-1)}{(\theta + \lambda x)} - x \right] = 0\,.$$

The score function

$$\frac{\partial}{\partial \theta_i} \ln P_x(\theta, \lambda)$$

is viewed as the relative rates of change in the probabilities as the parameters θ and λ change. This score function is being weighted by the relative frequency in the case of the MLE method and is weighted by the discrepancy between the relative frequency and the estimated probability in the case of the weighted discrepancies estimation method. Famoye and Lee (1992) considered the combination of these two methods to define the EWRC estimators. The EWRC estimating equations are

$$\sum_x f_x \,[f_x - P_x(\theta, \lambda)] \left[\frac{x(\theta + \lambda)}{\theta(\theta + \lambda x)} - 1 \right] = 0 \tag{9.44}$$

and

$$\sum_x f_x \,[f_x - P_x(\theta, \lambda)] \left[\frac{x(x-1)}{(\theta + \lambda x)} - x \right] = 0\,. \tag{9.45}$$

The bias under the EWRC estimation is as small or smaller than the bias from ML and moment estimation methods.

Lee and Famoye (1996) applied several methods to estimate the GPD parameters for fitting the number of chromosome aberrations under different dosages of radiations. They compared the methods of moments, ML, minimum chi-square, weighted discrepancy, and EWRC. They found that the EWRC method provided the smallest mean square error and mean absolute error for most dosages of radiation.

9.6.2 Interval Estimation

When the parameter λ is fixed at λ_0 in a small sample, Famoye and Consul (1990) showed that a $100(1-\alpha)\%$ CI (θ_ℓ, θ_u) for θ can be obtained by solving for θ_u and θ_ℓ in equations

$$\sum_{j=0}^{y} n\theta_u (n\theta_u + j\lambda_o)^{j-1}\, e^{-n\theta_u - j\lambda_o}/j\,! = \frac{\alpha}{2} \tag{9.46}$$

and

$$\sum_{j=y}^{\infty} n\theta_\ell (n\theta_\ell + j\lambda_o)^{j-1} e^{-n\theta_\ell - j\lambda_o}/j! = \frac{\alpha}{2}, \tag{9.47}$$

where y is the sample sum.

When the ML point estimate of θ is more than 10, a sharper $100(1-\alpha)\%$ CI for θ may be obtained by using the property of normal approximation. Thus, a $100(1-\alpha)\%$ CI for θ is given by

$$\frac{\bar{x}(1-\lambda_0)^3}{(1-\lambda_0)^2 + z_{\alpha/2}} < \theta < \frac{\bar{x}(1-\lambda_0)^3}{(1-\lambda_0)^2 - z_{\alpha/2}}. \tag{9.48}$$

For large sample size, a $100(1-\alpha)\%$ CI for θ is given by

$$\frac{\left(\bar{x} - z_{\alpha/2}s\right)(1-\lambda_0)}{\sqrt{n}} < \theta < \frac{\left(\bar{x} + z_{\alpha/2}s\right)(1-\lambda_0)}{\sqrt{n}}. \tag{9.49}$$

The statistic s in (9.49) may be dropped for $\sigma^2 = \theta(1-\lambda_0)^{-3}$. By using only the sample mean $\bar{x}$, a $100(1-\alpha)\%$ CI for θ becomes

$$\frac{\bar{x}(1-\lambda_0)^3\sqrt{n}}{(1-\lambda_0)^2\sqrt{n} + z_{\alpha/2}} < \theta < \frac{\bar{x}(1-\lambda_0)^3\sqrt{n}}{(1-\lambda_0)^2\sqrt{n} - z_{\alpha/2}}. \tag{9.50}$$

When the parameter θ is fixed at θ_0 in a small sample, a $100(1-\alpha)\%$ CI for λ can be obtained from equations (9.46) and (9.47) by replacing θ_u and θ_ℓ with θ_0 and by replacing λ_0 in (9.46) with λ_u and λ_0 in (9.47) with λ_ℓ. For large samples, a $100(1-\alpha)\%$ CI for λ, when statistics $\bar{x}$ and s are used, is given by

$$1 - \frac{\theta_0}{\bar{x} - z_{\alpha/2}s/\sqrt{n}} < \lambda < 1 - \frac{\theta_0}{\bar{x} + z_{\alpha/2}s/\sqrt{n}}. \tag{9.51}$$

If only the sample mean $\bar{x}$ is used, a $100(1-\alpha)\%$ CI for λ is given by finding the smallest value of λ that satisfies the inequality

$$[(1-\lambda)\,\bar{x} - \theta_0]\sqrt{n(1-\lambda)} + \sqrt{\theta_0}\, z_{\alpha/2} > 0, \tag{9.52}$$

and the largest value of λ that satisfies the inequality

$$[(1-\lambda)\,\bar{x} - \theta_0]\sqrt{n(1-\lambda)} - \sqrt{\theta_0}\, z_{\alpha/2} < 0. \tag{9.53}$$

The smallest and largest values of λ satisfying (9.52) and (9.53), respectively, may be determined with the help of a computer program, as given in the book by Consul (1989a).

Suppose the two parameters θ and λ are unknown and we wish to obtain CIs for one of the parameters. The parameter, which is not of interest, becomes a nuisance parameter and has to be eliminated before any inference can be made. The method of "maximization of likelihood" for eliminating the nuisance parameter can be applied. Let $\hat{\theta}$ and $\hat{\lambda}$ be the ML estimates of θ and λ, respectively. Famoye and Consul (1990) applied the method of "maximization of likelihood" and derived a U-shaped likelihood ratio statistic for determining a $100(1-\alpha)\%$ CI for θ when λ is a nuisance parameter. The statistic is given by

$$T_m(\theta) = -2n\left[\ln(\theta/\hat{\theta}) - \theta + \hat{\theta} - \bar{x}(\tilde{\lambda}(\theta) - \hat{\lambda})\right] - \sum_{i=1}^{n} 2(x_i - 1)\ln\left(\frac{\theta + \tilde{\lambda}(\theta)x_i}{\hat{\theta} + \hat{\lambda}x_i}\right). \tag{9.54}$$

A $100(1-\alpha)\%$ limits for θ are the values θ_ℓ and θ_u of θ at which the straight line

$$T_m(\theta) = \chi^2_{\alpha,1}$$

intersects the graph of the function $T_m(\theta)$ against θ. The value $\chi^2_{\alpha,1}$ is the upper percentage point of the chi-square distribution with 1 degree of freedom. The corresponding statistic for constructing a $100(1-\alpha)\%$ CI for parameter λ when θ is a nuisance parameter is given by

$$T_m(\lambda) = -2n\left[\ln(\tilde{\theta}(\lambda)/\hat{\theta}) + \hat{\theta} - \tilde{\theta}(\lambda) + \bar{x}(\hat{\lambda} - \lambda)\right] - \sum_{i=1}^{n} 2(x_i - 1)\,\ln\left(\frac{\tilde{\theta}(\lambda) + \lambda x_i}{\hat{\theta} + \hat{\lambda} x_i}\right). \tag{9.55}$$

9.6.3 Confidence Regions

The partial derivatives of the log likelihood function can be used to obtain an approximate chi-square expression for constructing confidence regions when the sample size is large. Famoye and Consul (1990) derived the bivariate log likelihood function

$$\begin{aligned} T(\theta,\lambda) = \frac{\theta(\theta+2\lambda)}{n(\theta-\theta\lambda+2\lambda)} &\left\{\left[\frac{n(1-\theta)}{\theta} + \sum_{i=1}^{n}\frac{x_i-1}{\theta+\lambda x_i}\right]^2 + \frac{1-\lambda}{2(\theta+2\lambda)}\right. \\ &\left.\times\left[n\theta - n\bar{x}(2-\lambda+2\lambda/\theta) + (1+2\lambda/\theta)\sum_{i=1}^{n}\frac{x_i(x_i-1)}{\theta+\lambda x_i}\right]^2\right\}, \end{aligned} \tag{9.56}$$

which has an asymptotic chi-square distribution with two degrees of freedom. An approximate $100(1-\alpha)\%$ confidence region for (θ,λ) is the set of values of θ and λ for which

$$T(\theta,\lambda) \le \chi^2_{\alpha,2}\,.$$

9.7 Statistical Testing

9.7.1 Test about Parameters

Consul and Shenton (1973a) showed that if the r.v. X has a GPD and if $\lambda < 0.5$, the standardized variate

$$z = \frac{X-\mu}{\sigma}$$

tends to a standard normal form as θ increases without limit. Accordingly, a test for

$$H_0:\ \lambda = \lambda_0 \le 0.5 \quad \text{against} \quad H_1:\ \lambda > \lambda_0$$

can be based on normal approximation. Famoye and Consul (1990) based the test on $\bar{X} = \frac{1}{n}\sum X_i$, and the critical region at a significance level α is $\bar{X} > C$, where

$$C = \theta_0(1-\lambda_0)^{-1} + z_\alpha\sqrt{\frac{\theta_0(1-\lambda_0)^{-3}}{n}}\,. \tag{9.57}$$

The power of the test is given by

$$\pi = 1 - \beta = P(\bar{X} > c \mid H_1)\,.$$

In large samples, the test for θ or λ can be based on the likelihood ratio test. For a test about θ, the likelihood ratio test statistic is

$$T = -2\left[n\hat{\theta} - n\theta_0 + y\hat{\lambda} - y\tilde{\lambda}(\theta_0) + n\,\ln(\theta_0/\hat{\theta}) + \sum_{i=1}^{n}(x_i - 1)\,\ln\left(\frac{\theta_0 + \tilde{\lambda}(\theta_0)x_i}{\hat{\theta} + \hat{\lambda}x_i}\right)\right]. \tag{9.58}$$

The null hypothesis $H_0 : \theta = \theta_0$ (against the alternative $H_1 : \theta \neq \theta_0$) is rejected if

$$T > \chi^2_{\alpha,1}\,.$$

To test the hypothesis $H_0 : \lambda = \lambda_0$ against an alternative composite hypothesis $H_1 : \lambda \neq \lambda_0$, a similar likelihood ratio test statistic as in (9.58) can be used.

Famoye and Consul (1990) obtained the power of the likelihood ratio test of $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$. The power is approximated by

$$\pi = 1 - \beta \simeq \int_a^\infty d\,\chi^2\left(1 + \frac{\gamma_1}{1 + 2\gamma_1}\right), \tag{9.59}$$

where $\chi^2(r)$ is a central chi-square variate with r degrees of freedom,

$$a = (1 + \gamma_1)(1 + 2\gamma_1)^{-1}\,\chi^2_{\alpha,1} \tag{9.60}$$

and

$$\gamma_1 = (\theta_1 - \theta_0)\,\frac{2n\left[\theta_1 - \theta_1\tilde{\lambda}(\theta_1) + 2\tilde{\lambda}(\theta_1)\right]}{\theta_1\left[\theta_1 + 2\tilde{\lambda}(\theta_1)\right]}\,. \tag{9.61}$$

In (9.61), θ_1 is the specified value of θ under the alternative hypothesis.

Fazal (1977) has considered an asymptotic test to decide whether there is an inequality between the mean and the variance in Poisson-like data. If the test indicates inequality between them, then the GPD is the appropriate model for the observed data.

9.7.2 Chi-Square Test

The goodness-of-fit test of the GPD can be based on the chi-square statistic

$$\chi^2 = \sum_{x=0}^{k}(O_x - E_x)^2/E_x, \tag{9.62}$$

where O_x and E_x are the observed and the expected frequencies for class x. The parameters θ and λ are estimated by the ML technique. The expected value E_x is computed by

$$E_x = nP_x(\theta, \lambda), \tag{9.63}$$

where n is the sample size.

The random variable χ^2 in (9.62) has an asymptotic chi-square distribution with $k - 1 - r$ degrees of freedom where r is the number of estimated parameters in the GPD.

9.7.3 Empirical Distribution Function Test

Let a random sample of size n be taken from the GPD model (9.1) and let n_x, $x = 0, 1, 2, \ldots, k$, be the observed frequencies for the different classes, where k is the largest of the observations.

An empirical distribution function (EDF) for the sample is defined as

$$F_n(x) = \frac{1}{n}\sum_{i=0}^{x} n_i, \quad x = 0, 1, 2, \ldots, k. \tag{9.64}$$

Let the GPD cdf be

$$F(x; \theta, \lambda) = \sum_{i=0}^{x} P_i(\theta, \lambda), \quad x \geq 0, \tag{9.65}$$

where $P_i(\theta, \lambda)$ is given by (9.1). The EDF statistics are those that measure the discrepancy between $F_n(x)$ in (9.64) and $F(x; \theta, \lambda)$ in (9.65). Let $\tilde{\theta}$ and $\tilde{\lambda}$ be the moment estimators of the GPD, based on the observed sample. To test the goodness-of-fit of the GPD, Famoye (1999) defined some EDF statistics analogous to the statistics defined for the continuous distributions.
(a) The Kolmogorov–Smirnov statistic

$$K_d = \sup_x \left| F_n(x) - F(x; \tilde{\theta}, \tilde{\lambda}) \right|. \tag{9.66}$$

(b) The modified Cramer–von Mises statistic

$$W_d^* = n \sum_{x=0}^{k} \left[F_n(x) - F(x; \tilde{\theta}, \tilde{\lambda}) \right]^2 P_x(\tilde{\theta}, \tilde{\lambda}).$$

(c) The modified Anderson–Darling statistic

$$A_d^* = n \sum_{x=0}^{k} \frac{\left[F_n(x) - F(x; \tilde{\theta}, \tilde{\lambda}) \right]^2 P_x(\tilde{\theta}, \tilde{\lambda})}{F(x; \tilde{\theta}, \tilde{\lambda}) \left[1 - F(x; \tilde{\theta}, \tilde{\lambda}) \right]}.$$

The EDF test statistics use more information in the data than the chi-square goodness-of-fit test. By using the parametric bootstrap method, Famoye (1999) has carried out Monte Carlo simulations to estimate the critical values of the above three EDF test statistics and has shown that, in general, the Anderson–Darlin statistic A_d^* is the most powerful of all the EDF test statistics for testing the goodness-of-fit of the GPD model.

9.8 Characterizations

A large number of characteristic properties of the GPD have been provided in section 9.1 through section 9.5. Ahsanullah (1991a) used the property of infinite divisibility to characterize the GPD. We next consider some general probabilistic and statistical properties which lead to different characterizations of the GPD. For the proofs of these characterization theorems, the reader is referred to the book by Consul (1989a).

The following characterizations are based on the conditional probability.

Theorem 9.3. *Let X_1 and X_2 be two independent discrete r.v.s whose sum Z is a GP variate with parameters θ and λ as defined in (9.1). Then X_1 and X_2 must each be a GP variate defined over all nonnegative integers (Consul, 1974).*

Theorem 9.4. *If X_1 and X_2 are two independent GP variates with parameters (θ_1, λ) and (θ_2, λ), respectively, then the conditional probability distribution of X_1 given $X_1 + X_2 = n$ is a QBD-II (Consul, 1975).*

Theorem 9.5. *If a nonnegative GP variate N is subdivided into two components X and Y in such a way that the conditional distribution $P(X = k,\ Y = n - k \mid N = n)$ is QBD-II with parameters (n, p, θ), then the random variables X and Y are independent and have GP distributions (Consul, 1974).*

Theorem 9.6. *If X and Y are two independent r.v.s defined on the set of all nonnegative integers such that*

$$P(X = k \mid\ X + Y = n) = \frac{\binom{n}{k} p_n \pi_n \,(p_n + k\lambda)^{k-1}\, [\pi_n + (n-k)\lambda]^{n-k-1}}{(1 + n\lambda)^{n-1}} \tag{9.67}$$

for $k = 0, 1, 2, \ldots, n$, and zero otherwise, where $p_n + \pi_n = 1$, then

(a) *p_n is independent of n and equals a constant p for all values of n, and*
(b) *X and Y must have GP distributions with parameters $(p\alpha, \lambda\alpha)$ and $(\pi\alpha, \lambda\alpha)$, respectively, where $\alpha (> 0)$ is an arbitrary number (Consul, 1974).*

Theorem 9.7. *If X and Y are two independent nonnegative integer-valued r.v.s such that*

(a)
$$P(Y = 0 \mid X + Y = n) = \frac{\theta_1(\theta_1 + n\alpha)^{n-1}}{(\theta_1 + \theta_2)(\theta_1 + \theta_2 + n\alpha)^{n-1}}, \tag{9.68}$$

(b)
$$P(Y = 1 \mid X + Y = n) = \frac{n\theta_1\theta_2(\theta_1 + n\alpha - \alpha)^{n-2}}{(\theta_1 + \theta_2)(\theta_1 + \theta_2 + n\alpha)^{n-1}}, \tag{9.69}$$

where $\theta_1 > 0,\ \theta_2 > 0,\ 0 \le \alpha \le 1$. Then X and Y are GP variates with parameters $(\theta_1 p, \alpha p)$ and $(\theta_2 p, \alpha p)$, respectively, where p is an arbitrary number $0 < p < 1$ (Consul, 1975).

Situations often arise where the original observations produced by nature undergo a destructive process and what is recorded is only the damaged portion of the actual happenings. Consul (1975) stated and proved the following three theorems on characterizations by damage process.

Theorem 9.8. *If N is a GP variate with parameters $(\theta, \alpha\theta)$ and if the destructive process is QBD-II, given by*

$$S(k \mid n) = \binom{n}{k} \frac{p\pi}{1 + n\alpha} \left(\frac{p + k\alpha}{1 + n\alpha}\right)^{k-1} \left(\frac{\pi + (n-k)\alpha}{1 + n\alpha}\right)^{n-k-1}, \quad k = 0, 1, 2, \ldots, n, \tag{9.70}$$

where $0 < p < 1,\ p + \pi = 1,\ \alpha > 0$, and Y is the undamaged part of N, then

(a) *Y is a GP variate with parameters $(p\theta, \alpha\theta)$,*

(b) $P(Y = k) = P(Y = k \mid N\ damaged) = P(Y = k \mid N\ undamaged),$ *and*
(c) $S_k = 0$ *for all* k *if* $S_k = P(Y = k) - P(Y = k \mid N\ undamaged)$ *does not change its sign for any integral value of* k.

Theorem 9.9. *Suppose that* $S(k \mid n)$ *denotes the QBD-II given by (9.70). Then* $P(Y = k) = P(Y = k \mid X\ undamaged)$ *if and only if* $\{P_x\}$ *is a GP variate.*

Theorem 9.10. *If a GP variate* N, *with parameters* $(\theta, \theta\alpha)$, *gets damaged by a destructive process* $S(k \mid n)$ *and is reduced to a variate* Y *such that*

$$P(Y = k) = P(Y = k \mid N\ undamaged),$$

the destructive process $S(k \mid n)$ *must be QBD-II.*

Theorem 9.11. *Let* $X_1, X_2, \ldots, X_N$ *be a random sample taken from a discrete population possessing the first three moments. Let*

$$\Lambda = X_1 + X_2 + \cdots + X_N ,$$

and let a statistic T *be defined in terms of the eight subscripts* $g, h, \ldots, n$, *by*

$$\begin{aligned} T &= 120 \sum X_g X_h \cdots X_m \, [28X_n + (N-7)(14 - 3X_m)] \\ &\quad - 20(N-6)^{(2)} \sum X_g \cdots X_k X_\ell^2 \, [X_\ell - 3X_k + 6] \\ &\quad + 6(N-5)^{(3)} \sum X_g X_h X_i X_j^2 X_k^2 \, [2X_k + 2 - 3X_i] \\ &\quad - (N-4)^{(4)} \sum X_g X_h \cdots X_i^2 X_j^2 \left[X_n X_j - \frac{1}{3} X_i X_j + 2X_n - 18 X_g X_n \right], \end{aligned} \tag{9.71}$$

where $(N)^{(j)} = N!/(N-j)!$ *and the summations are taken over all subscripts* $g, h, i, \ldots, n$ *which are different from each other and vary from 1 to N. Then the population must be a GPD if and only if the statistic* T *has a zero regression on the statistic* Λ *(Consul and Gupta, 1975).*

9.9 Applications

Chromosomes are damaged one or more times in the production process, and zero or more damages are repaired in the restitution process. The undamaged chromosomes form a queue in the production process and the damaged ones form a queue in the restitution process. Janardan and Schaeffer (1977) have shown that if X is the net number of aberrations (damaged chromosomes) awaiting restitution in the queue, then the probability distribution of the r.v. X is the GPD given by (9.1). It was suggested that the parameter λ in the GPD represented an equilibrium constant which is the limit of the ratio of the rate of induction to the rate of restitution, and thus the GPD could be used to estimate the net free energy for the production of induced chromosome aberrations.

Consul (1989a) described the application of GPD to the number of chromosome aberrations induced by chemical and physical agents in human and animal cells. Janardan, Schaeffer, and DuFrain (1981) observed that a three-parameter infinite mixture of Poisson distributions is slightly better than the GPD when the parameters are estimated by the moment method. Consul (1989a) used the ML estimates and found that the fit by the GPD model was extremely good.

Schaeffer et al. (1983) provided a formal link between the GPD and the thermodynamic free energy by using a Markov-chain model and estimated that the free energy required to produce isochromatid breaks or dicentrics is about 3.67 KJ/mole/aberration and 18.4 KJ/mole, which is in good agreement with free energy estimates on the formation of DNA. A detailed description of this model can be studied either in their paper or in Consul (1989a), where many other models and applications are also given.

Janardan, Kerster, and Schaeffer (1979) considered sets of data on spiders and sow bugs, weevil eggs per bean, and data on sea urchin eggs. They showed that the observed patterns can be easily explained and described by the GPD models. Interpretative meanings were given to the parameters θ and λ in GPD.

Consul (1989a) also described the use of GPD to study shunting accidents, home injuries, and strikes in industries. Meanings were given to both parameters θ and λ in the applications.

The number of units of different commodities purchased by consumers over a period of time follows the GPD model. Consul (1989a) suggested the following interpretations for the parameter values: θ reflects the basic sales potential for the product and λ represents the average rates of liking generated by the product among consumers.

Other important applications discussed by Consul (1989a) are references of authors, spatial patterns, diffusion of information, and traffic and vehicle occupancy.

Tripathi, Gupta, and Gupta (1986) have given a very interesting use of the GPD in the textile manufacturing industry. Since the Poisson distribution, a particular case of the GPD, is generally used in the industry, they compared it with the GPD for increasing the profits. Let X be a random variable which represents the characteristic of a certain product and let x be its observed value. The profit $P(x)$ equals the amount received for good product plus the amount received for scrap (unusable) product minus the manufacturing cost of the total product and minus the fixed cost. They took $\mathrm{E}[P(X)]$ and obtained the condition for its maximization. By considering different values for the various parameters in the problem they found that in each case the profits were larger when the GPD was used instead of the Poisson distribution.

Itoh, Inagaki, and Saslaw (1993) showed that when clusters are from Poisson initial conditions, the evolved Eulerian distribution is generalized Poisson. The GPD provides a good fit to the distribution of particle counts in randomly placed cells, provided the particle distributions evolved as a result of gravitational clustering from an initial Poisson distribution. In an application in astrophysics, Sheth (1998) presented a derivation of the GPD based on the barrier crossing statistics of random walks associated with Poisson distribution.

9.10 Truncated Generalized Poisson Distribution

Consul and Famoye (1989b) defined the truncated GPD by

$$\Pr(X = x) = P_x(\theta, \lambda)/F_m(\theta, \lambda), \qquad x = 0, 1, 2, \ldots, m, \tag{9.72}$$

and zero otherwise, where $\theta > 0$, $-\infty < \lambda < \infty$ and $F_m(\theta, \lambda)$ is given by (9.2). In (9.72), m is any positive integer less than or equal to the largest possible value of x such that $\theta + \lambda x > 0$.

When $\lambda < -\theta/2$, the truncated GPD reduces to the point binomial model with $m = 1$ and probabilities

$$P(X = 0) = \left(1 + \theta e^{-\lambda}\right)^{-1} \tag{9.73}$$

and

$$P(X=1) = \theta e^{-\lambda} \left(1+\theta e^{-\lambda}\right)^{-1} . \tag{9.74}$$

When $\lambda > -\theta/2$, the value of m can be any positive integer $\leq -\theta/\lambda$. When $0 < \lambda < 1$, the largest value of m is $+\infty$ and the truncated GPD reduces to the GPD model (9.1) since $F_m(\theta, \lambda) = 1$ for $m = \infty$. When $\lambda > 1$, the quantity $P_x(\theta, \lambda)$ is positive for all integral values of x and the largest value of m is $+\infty$. However, $F_\infty(\theta, \lambda)$ is not unity.

The mean μ_m and variance σ_m^2 of the truncated GPD can be written in the form

$$\mu_m = E(X) = [F_m(\theta, \lambda)]^{-1} \sum_{x=1}^{m} x P_x(\theta, \lambda) \tag{9.75}$$

and

$$\sigma_m^2 = [F_m(\theta, \lambda)]^{-1} \sum_{x=1}^{m} x^2 P_x(\theta, \lambda) - \mu_m^2 . \tag{9.76}$$

Consul and Famoye (1989b) considered the ML estimation of the parameters of truncated GPD. They also obtained estimates based upon the mean and ratio of the first two frequencies.

Shanmugam (1984) took a random sample $X_i, \ \ i = 1, 2, \ldots, n$, from a positive (zero truncated or decapitated) GPD given by

$$P(X=x) = \lambda(1+\alpha x)^{x-1} \left(\theta e^{-\alpha\theta}\right)^x \Big/ x!, \quad x = 1, 2, 3, \ldots,$$

where $\lambda = \left(e^\theta - 1\right)^{-1}$, and obtained a statistic, based on the sample sum $\sum_{i=1}^n X_i = k$ to test the homogeneity of the random sample.

The probability distribution of the sample sum is given by

$$P\left(\sum_{i=1}^n X_i = k\right) = \lambda^{-n} n!\, t(k, n, \alpha) \left(\theta e^{-\alpha\theta}\right)^k \Big/ k!$$

for $k = n, n+1, n+2, \ldots$ and where

$$t(k, n, \alpha) = \sum_{i=0}^{k-1} \binom{k-1}{i} (\alpha k)^{k-i-1} S(n, i+1),$$

$S(n, i+1)$ being the Stirling numbers of the second kind. The conditional distribution of X_1, given $\sum_{i=1}^n X_i = k$, is then

$$P\left(X_1 = x \,\middle|\, \sum_{i=1}^n X_i = k\right) = \frac{\binom{k}{x}(1+\alpha x)^{x-1} t(k-x, n-1, \alpha)}{n\, t(k, n, \alpha)}, \quad x = 1, 2, \ldots, k-n+1.$$

Since the positive GPD is a modified power series distribution and the above expression is independent of θ, the sum $\sum_{i=1}^n X_i$ becomes a complete sufficient statistic for θ and the above expression provides an MVU estimate for the probability of the positive GPD. The above conditional distribution provides a characterization also that the mutually independent positive integer-valued r.v.s $X_1, X_2, \ldots, X_n, \ \ n \geq 2$, have the same positive GPD.

Shanmugam (1984) defines a statistic

$$V = \sum_{i=1}^{n} X_i^2 \quad \text{for fixed} \quad \sum_{i=1}^{n} X_i = k,$$

and shows that though the r.v.s $X_1, X_2, \ldots, X_n$, are mutually dependent on account of the fixed sum, yet they are asymptotically mutually independent and V is asymptotically normally distributed. Under the null hypothesis

$$H_0 : (X_1, X_2, \ldots, X_n) \quad \text{is a homogeneous random sample from positive GPD,}$$

Shanmugam obtains complex expressions for

$$\mu_0 = \mathrm{E}\left[V \left| \sum_{i=1}^{n} X_i \right. = k \right], \quad \sigma_0^2 = \mathrm{Var}\left[V \left| \sum_{i=1}^{n} X_i \right. = k \right].$$

The null hypothesis H_0 is rejected if

$$\sigma_0^{-1} \mid V - \mu_0 \mid \geq z_{\epsilon/2},$$

where $z_{\epsilon/2}$ is the $(1 - \epsilon/2)$th percentile of the unit normal distribution.

9.11 Restricted Generalized Poisson Distribution

9.11.1 Introduction and Definition

In many applied problems, it is known in advance that the second parameter λ in the GPD model (9.1) is linearly proportional to the parameter θ (see an example in Consul, 1989a, section 2.6). Putting $\lambda = \alpha\theta$ in the model (9.1), we get the probabilities for the restricted GPD model in the form

$$P_x(\theta, \alpha\theta) = (1 + x\alpha)^{x-1} \theta^x e^{-\theta - x\alpha\theta}/x!, \quad x = 0, 1, 2, \ldots, \tag{9.77}$$

and zero otherwise. For the probability model in (9.77), the domain of α is $\max(-\theta^{-1}, -m^{-1}) < \alpha < \theta^{-1}$, and accordingly, the parameter α is restricted above by θ^{-1}. In (9.77), $P_x(\theta, \alpha\theta) = 0$ for $x > m$ when $\alpha < 0$. The model reduces to the Poisson distribution when $\alpha = 0$.

The mean and variance of the restricted GPD are given by

$$\mu = \theta(1 - \alpha\theta)^{-1} \quad \text{and} \quad \sigma^2 = \theta(1 - \alpha\theta)^{-3}. \tag{9.78}$$

Other higher moments can be obtained from (9.13) by replacing λ with $\alpha\theta$.

9.11.2 Estimation

The restricted GPD is a MPSD defined and discussed in chapter 7 with

$$\phi(\theta) = \theta\, e^{-\alpha\theta} \quad \text{and} \quad h(\theta) = e^{\theta}.$$

When α is known, all the results derived for the MPSD in chapter 7 hold for the restricted GPD. Thus, the ML estimation of θ, MVU estimation of θ, and its function $\ell(\theta)$ for both restricted GPD and truncated restricted GPD are similar to the results obtained for the MPSD in chapter 7.

When both parameters θ and α are unknown in (9.77), the moment estimators are

$$\tilde{\theta} = \sqrt{\frac{\bar{x}^3}{s^2}} \quad \text{and} \quad \tilde{\alpha} = \sqrt{\frac{s^2}{\bar{x}^3}} - \frac{1}{\bar{x}} . \tag{9.79}$$

Kumar and Consul (1980) obtained the asymptotic biases, variances, and covariance of the moment estimators in (9.79) as

$$b(\tilde{\theta}) \simeq \frac{\theta}{4n} \left[5 + \frac{3\alpha(2 + 3\alpha\theta)}{1 - \alpha\theta} \right] , \tag{9.80}$$

$$b(\tilde{\alpha}) \simeq \frac{-3}{4n\theta} \left[1 + \frac{2\alpha + \alpha^2\theta}{1 - \alpha\theta} \right] , \tag{9.81}$$

$$V(\tilde{\theta}) \simeq \frac{\theta}{2n} \left[\theta + \frac{2 - 2\alpha\theta + 3\alpha^2\theta^2}{1 - \alpha\theta} \right] , \tag{9.82}$$

$$V(\tilde{\alpha}) \simeq \frac{1}{2n\theta^2} \left[1 + \frac{2\alpha + \alpha^2\theta}{1 - \alpha\theta} \right] , \tag{9.83}$$

and

$$Cov(\tilde{\theta}, \tilde{\alpha}) \simeq \frac{-1}{2n} \left[1 + \frac{2\alpha + \alpha^2\theta}{1 - \alpha\theta} \right] . \tag{9.84}$$

The generalized variance of $\tilde{\theta}$ and $\tilde{\alpha}$ is given by

$$\left| \sum \right| \simeq \frac{n-1}{2n^3\theta} \left[1 - \alpha\theta + \frac{n-1}{n} \left(2\alpha + \alpha^2\theta \right) \right] . \tag{9.85}$$

For estimators based on sample mean and zero-class frequency, Consul (1989a) gave the following:

$$\theta^* = -\ln\left(\frac{n_0}{n}\right) \quad \text{and} \quad a^* = (\theta^*)^{-1} - \frac{1}{\bar{x}} . \tag{9.86}$$

The variance of θ^* is the same as given in (9.33). However, the variance of α^* and the covariance of θ^* and α^* are given by

$$V(\alpha^*) \simeq \frac{1}{n\theta^4} \left[e^\theta - 1 - \theta(1 - \alpha\theta) \right] \tag{9.87}$$

and

$$Cov(\theta^*, \alpha^*) \simeq \frac{-1}{n\theta^2} \left[e^\theta - 1 - \theta(1 - \alpha\theta) \right] . \tag{9.88}$$

The ML estimate $\hat{\theta}$ of θ in restricted GPD is found by solving

$$\sum_{i=1}^{n} \frac{x_i(x_i - 1)}{\theta(\bar{x} - x_i) + x_i\bar{x}} - n = 0 \tag{9.89}$$

iteratively. The corresponding ML estimate $\hat{\alpha}$ of α is obtained from

$$\hat{\alpha} = \hat{\theta}^{-1} - \frac{1}{\bar{x}} .$$

The uniqueness of the ML estimates for the restricted GPD has not been shown. Consul (1989a) conjectured that the estimate will be unique and that $\hat{\theta}$ will be greater than $\bar{x}$ or less than $\bar{x}$ according to whether the sample variance is greater or less than the, sample mean.

The asymptotic biases, variances, and covariance are

$$b(\tilde{\theta}) \simeq \frac{\theta(5+12\alpha)}{4n(1+3\alpha)} , \tag{9.90}$$

$$b(\hat{\alpha}) \simeq \frac{-3}{4n\theta}\,\frac{1-2\alpha}{1+3\alpha} , \tag{9.91}$$

$$V(\hat{\theta}) \simeq \frac{\theta(2+\theta)}{2n} , \tag{9.92}$$

$$V(\hat{\alpha}) \simeq \frac{1+2\alpha}{2n\theta^2} , \tag{9.93}$$

and

$$Cov(\hat{\theta}, \hat{\alpha}) \simeq \frac{-(1+2\alpha)}{2n} . \tag{9.94}$$

Let $\Phi_1 = \ln\theta$ and $\Phi_2 = \ln(1-\alpha\theta)$, so that $\Phi' = (\Phi_1, \Phi_2)$ denote the new parameters for estimation and

$$\left.\begin{aligned} \eta_1 &= \ln\mu = \ln\theta - \ln(1-\alpha\theta) = \Phi_1 - \Phi_2 \\ \eta_2 &= \ln\mu_2 = \ln\theta - 3\ln(1-\alpha\theta) = \Phi_1 - 3\Phi_2 \\ \eta_3 &= \ln(-\ln P_0) = \ln\theta = \Phi_1 . \end{aligned}\right\} \tag{9.95}$$

We note that $\ln(\bar{x})$, $\ln(s^2)$, and $\ln[-\ln(n_0/n)]$ are the sample estimates for η_1, η_2, and η_3, respectively.

By using

$$h' = \Big(\ln(\bar{x}),\ \ln(s^2),\ \ln[-\ln(n_0/n)]\Big)$$

and

$$\eta' = \big(\eta_1, \eta_2, \eta_3\big) \quad \text{with} \quad \eta = W\ \Phi$$

where

$$W = \begin{bmatrix} 1 & -1 \\ 1 & -3 \\ 1 & 0 \end{bmatrix},$$

Consul (1989a) derived a generalized minimum chi-square estimators $\tilde{\theta}$ and $\tilde{\alpha}$ for θ or α. The estimators are given by

$$\tilde{\theta} = \ \exp(\Phi_1) \quad \text{and} \quad \tilde{\alpha} = \big(1 - e^{\Phi_2}\big)/\tilde{\theta}. \tag{9.96}$$

When the sample is small and parameter α in restricted GPD is known, a $100(1-\alpha)\%$ CI for θ can be based on the statistic Y which is complete and sufficient for θ. The result is similar to the interval estimation in section 9.6 for the GPD model in (9.1).

When the two parameters θ and α are unknown in the model (9.77), the method of maximization of likelihood can be used to eliminate the nuisance parameter. In addition to this method, Famoye and Consul (1990) proposed the method of conditioning for eliminating parameter θ in the restricted GPD. This approach led to the statistic

$$T_c(\alpha) = -2\left[(n\bar{x}-1)\ln\left(\frac{1-\hat{\alpha}}{1+\alpha\bar{x}}\right) + \sum_{l=1}^{n}(x_i-1)\ln\left(\frac{1+\alpha x_i}{1+2x_i}\right)\right]. \tag{9.97}$$

The function $T_c(\alpha)$ is U-shaped and a $100(1-\alpha)\%$ CI for α can be obtained by finding the two values α_ℓ and α_u at which the straight line

$$T_c(\alpha) = \chi^2_{\alpha,1}$$

intersects the graph of $T_c(\alpha)$ in (9.97).

9.11.3 Hypothesis Testing

All the tests described in section 9.7 of this chapter are applicable to the restricted GPD model. In addition, a uniformly most powerful test for θ when α is known can be constructed. Famoye and Consul (1990) described a uniformly most powerful test for testing

$$H_0 : \theta \le \theta_0 \quad \text{against} \quad H_1 : \theta > \theta_0.$$

Consider the null hypothesis

$$H_0 : \theta = \theta_0 \quad \text{against} \quad H_1 : \theta = \theta_1 \;\; (\theta_1 > \theta_0). \tag{9.98}$$

If $X_1, X_2, \ldots$ is a sequence of independent r.v.s from the restricted GPD in (9.77) and the value of α is known, Consul (1989a) proposed a sequential probability ratio test for the hypotheses in (9.98). The test is to observe $\{X_i\}$, $i = 1, 2, \ldots, N$, successively and at any state $N \ge 1$,

(i) reject H_0 if $L(\underline{x}) \ge A$,
(ii) accept H_0 if $L(\underline{x}) \le B$,
(iii) continue observing X_{N+1} if $B < L(\underline{x}) < A$, where

$$L(\underline{x}) = \frac{\theta_1}{\theta_0}\sum_{i=1}^{N} x_i \; e^{\left(N+\alpha\sum_{i=1}^{N} x_i\right)(\theta_0-\theta_1)}. \tag{9.99}$$

The constants A and B are approximated by

$$A \simeq \frac{1-\beta_1}{\alpha_1} \quad \text{and} \quad B \simeq \frac{\beta_1}{1-\alpha_1}, \tag{9.100}$$

where α_1 and β_1 are the probabilities of type I and type II errors, respectively.

Let Z_i, $i = 1, 2, 3, \ldots, r$, $r+1$ be independent restricted GP variates with parameters (θ_i, α_i). Famoye (1993) developed test statistics to test the homogeneity hypothesis

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_{r+1} = \theta \tag{9.101}$$

against a general class of alternatives. When θ is known, the test statistic is

$$T = (1-\alpha\theta)^2 \sum_{i=1}^{r+1} \left(Z_i - \frac{\theta}{1-\alpha\theta} \right)^2 - \sum_{i=1}^{r+1} Z_i, \tag{9.102}$$

which can be approximated by a normal distribution. The mean and variance of T are

$$\mathrm{E}(T) = 0 \quad \text{and} \quad \mathrm{Var}(T) = \frac{2\theta^2 (r+1)(1-\alpha\theta+2\alpha+3\alpha^2\theta)}{(1-\alpha\theta)^3}. \tag{9.103}$$

When θ is unknown, a test of homogeneity for the restricted GPD against a general class of alternatives is based on a large value of $\sum_{i=1}^{r+1} Z_i^2$ conditional on the sample sum $\sum_{i=1}^{r+1} Z_i = m$.

9.12 Other Related Distributions

9.12.1 Compound and Weighted GPD

Goovaerts and Kaas (1991) defined a random variable S by

$$S = X_1 + X_2 + \cdots + X_N,$$

where $X_i,\ i = 1, 2, \ldots, N$, denote the amounts of the claims under the different insurance policies and N is the number of claims produced by a portfolio of policies in a given time period. Assuming that the random variables $N, X_1, X_2, \ldots$ are mutually independent, that $X_1, X_2, \ldots, X_N$ are identically distributed r.v.s with the distribution function $F(x)$, and that N has the GPD, they obtained the distribution function of S as

$$F_S(x) = \sum_{n=0}^{\infty} F^{*n}(x)\lambda(\lambda+n\theta)^{n-1} e^{-\lambda-n\theta}/(n!)$$

and called it the compound generalized Poisson distribution (CGPD). They used a recursive method, involving Panjer's recursion, to compute the total claims distribution of S.

Ambagaspitiya and Balakrishnan (1994) have obtained the pgf $P_S(z)$ and the mgf $M_S(z)$ of the CGPD, the total claim amount distribution in terms of the Lambert's W function as

$$P_S(z) = \exp\left\{-(\lambda/\theta)\left[W(-\theta\exp(-\theta)P_X(z)) + \theta\right]\right\}$$

and

$$M_S(z) = \exp\left\{-(\lambda/\theta)\left[W(-\theta\exp(-\theta)M_X(z)) + \theta\right]\right\},$$

where $P_X(z),\ M_X(z)$, and W are defined in section 9.2. They have also derived an integral equation for the probability distribution function of CGPD, when the insurance claim severities are absolutely continuous and have given a recursive formula for the probability function of CGPD.

By using Rao's (1965) definition of a weighted distribution

$$f_x(\theta,\lambda) = \frac{\omega(\theta+\lambda x,\lambda)P_x(\theta,\lambda)}{W(\theta,\lambda) = \sum_x \omega(\theta+\lambda x,\lambda)P_x(\theta,\lambda)},$$

Ambagaspitiya (1995) has defined a weighted GPD as

$$P_x(\theta, \lambda) = \frac{W(\theta+\lambda, \lambda)}{W(\theta, \lambda)} \frac{\theta}{\theta+\lambda} \left(\lambda + \frac{\theta}{x}\right) P_{x-1}(\theta+\lambda, \lambda)$$

for $x = 1, 2, 3, \ldots$. When $\omega(\theta, \lambda) = \theta$, the above reduces to the linear function Poisson distribution (Jain, 1975a). He obtained the pgf of the weighted GPD above and showed that it satisfies the convolution property.

Also, Ambagaspitiya (1995) considered a discrete distribution family with the property

$$P_x(a, b) = \left[h_1(a, b) + x^{-1} h_2(a, b)\right] P_{x-1}(a+b, b), \quad x = 1, 2, 3, \ldots,$$

and showed that the weighted GPD, with weight of the form $\omega(\theta + \lambda x, \lambda)$, forms a subclass of this family. A recursive formula for computing the distribution function has been provided.

9.12.2 Differences of Two GP Variates

Suppose that X has the distribution $P_x(\theta_1, \lambda)$ and Y has the distribution $P_y(\theta_2, \lambda)$ and X and Y are independent. Consul (1986) showed that the probability distribution of $D = X - Y$ is

$$P(D = X - Y = d) = e^{-\theta_1 - \theta_2 - d\lambda} \sum_{y=0}^{\infty} (\theta_1, \lambda)_{y+d} (\theta_2, \lambda)_y \, e^{-2y\lambda}, \tag{9.104}$$

where

$$(\theta, \lambda)_x = \frac{\theta(\theta + x\lambda)^{x-1}}{x!} \tag{9.105}$$

and d takes all integral values from $-\infty$ to $+\infty$.

The pgf of the r.v. $D = X - Y$ is

$$G(u) = \exp\left[\theta_1(z_1 - 1) + \theta_2(z_2 - 1)\right], \tag{9.106}$$

where $z_1 = u\, e^{\lambda(z_1 - 1)}$ and $z_2 = u^{-1}\, e^{\lambda(z_2 - 1)}$.

From (9.106), the cgf of D is

$$\psi(\beta) = \frac{\theta_1(Z_1 - \beta)}{\lambda} + \frac{\theta_2(Z_2 - \beta)}{\lambda}, \tag{9.107}$$

where $Z_1 = \beta + \lambda\left(e^{Z_1} - 1\right)$ and $Z_2 = -\beta + \lambda\left(e^{Z_2} - 1\right)$.

Consul (1989a) denoted the cumulants by L_k, $k = 1, 2, 3, \ldots$, and obtained the relation

$$(1-\lambda)\, L_{k+1} = \left(1 + \lambda \frac{\partial}{\partial \lambda}\right) \left(2\theta_1 \frac{\partial}{\partial \theta_1} - 1\right) L_k, \qquad k = 1, 2, 3, \ldots. \tag{9.108}$$

The first four cumulants are given by

$$L_1 = \frac{\theta_1 - \theta_2}{1 - \lambda}, \tag{9.109}$$

$$L_2 = \frac{\theta_1 + \theta_2}{(1-\lambda)^3}, \tag{9.110}$$

$$L_3 = \frac{(\theta_1 - \theta_2)(1 + 2\lambda)}{(1-\lambda)^5}, \tag{9.111}$$

and

$$L_4 = \frac{(\theta_1 + \theta_2)(1 + 8\lambda + 6\lambda^2)}{(1-\lambda)^7}. \tag{9.112}$$

The coefficients of skewness and kurtosis for the r.v. D are

$$\beta_1 = \frac{(\theta_1 - \theta_2)^2}{(\theta_1 + \theta_2)^3} \frac{(1+2\lambda)^2}{1-\lambda} \quad \text{and} \quad \beta_2 = 3 + \frac{1 + 8\lambda + 6\lambda^2}{(\theta_1 + \theta_2)(1-\lambda)}. \tag{9.113}$$

9.12.3 Absolute Difference of Two GP Variates

Let X_1 and X_2 be two independent GP variates and let $Y = |X_1 - X_2|$ be the absolute difference. Let the probability distributions of the two variates be given by

$$\begin{aligned} P_i^{(j)} &= \frac{(1+i\lambda)^{i-1}}{i!} \left(\theta_j\, e^{-\lambda\theta_j}\right)^i\, e^{-\theta_j}, \qquad i = 0, 1, 2, \ldots, \quad j = 1, 2, \\ &= C(i, \lambda)(\varphi_j)^i\, e^{-\theta_j} \end{aligned} \tag{9.114}$$

and zero otherwise, where $C(i, \lambda) = (1+i\lambda)^{i-1}/i!$ and $\varphi_j = \theta_j\, e^{-\lambda\theta_j}$. By using

$$\varphi_j' = \frac{d\varphi_j}{d\theta_j} = (1 - \lambda\theta_j)\, e^{-\lambda\theta_j},$$

Consul (1986) showed that the probability distribution of the r.v. Y is given by

$$P(Y = k) = \begin{cases} \sum_{i=0}^{\infty} [C(i,\lambda)]^2\, (\varphi_1 . \varphi_2)^i\, e^{-\theta_1 - \theta_2}, & k = 0, \\ \sum_{i=0}^{\infty} C(i,\lambda) C(i+k,\lambda)\, e^{-\theta_1-\theta_2} \left[\varphi_1^{i+k}\, \varphi_2^i + \varphi_1^i \varphi_2^{i+k}\right], & k = 1, 2, 3, \ldots. \end{cases} \tag{9.115}$$

9.12.4 Distribution of Order Statistics when Sample Size Is a GP Variate

Suppose $X_i,\ i = 1, 2, \ldots, N$, is a random sample of size N from a population with pdf $f(x)$ and cdf $F(x)$. Suppose $Y_1, Y_2, \ldots, Y_N$ denote the corresponding order statistics and the sample size N is a restricted GP variate with probability distribution $P_x(\theta, \alpha\theta)$ in (9.77).

Let $g_j(y \mid n)$ denote the conditional pdf of the jth order statistics Y_j for a given $N = n$, let h_j be the unconditional pdf of Y_j, and let h_{ij} be the joint unconditional pdf of Y_i and Y_j. Consul (1984) showed that the unconditional pdf of the jth order statistics Y_j is given by

$$h_j = \frac{\left(\theta\, e^{-\alpha\theta}\right)^j\, F_j^{j-1}\, f_j}{(j-1)!\, e^{\theta}\, Q_j(\theta, \alpha\theta)} \sum_{r=0}^{\infty} \frac{(1 + \alpha j + \alpha r)^{r+j-1}}{r!} \left[\theta(1 - F_j)\, e^{-\alpha\theta}\right]^r, \tag{9.116}$$

where

$$Q_j(\theta, \alpha\theta) = \sum_{i=j+1}^{\infty} P_i(\theta, \alpha\theta). \tag{9.117}$$

Also, the joint pdf of Y_i and $Y_j,\ i < j$, is given by

$$h_{ij} = \frac{F_j^{i-1}\, \left(F_j - F_i\right)^{j-i-1} f_i\, f_j \left(\theta\, e^{-\alpha\theta}\right)^j}{(i-1)!(j-i-1)!\, e^{\theta}\, Q_j(\theta, \alpha\theta)} \sum_{r=0}^{\infty} \frac{(1 + \alpha j + \alpha r)^{r+j-1}}{r!} \left(1 - F_j\right)^r\, \theta^r\, e^{-r\alpha\theta}, \tag{9.118}$$

where F_i and F_j are the distribution functions of Y_i and Y_j, respectively.

9.12.5 The Normal and Inverse Gaussian Distributions

Let X be a GP variate with parameters θ and λ. Consul and Shenton (1973a) showed that for all values of λ, the random variable

$$Z = \frac{X - \mu}{\sigma} \tag{9.119}$$

approaches the standard normal curve as θ becomes infinitely large. When $-0.5 < \lambda < 0.2$, a value of θ such as 15 makes the GPD model approximately normal.

If X is a GP variate with mean μ and variance σ^2, Consul and Shenton (1973a) showed that the distribution of

$$Y = X/\sigma$$

approaches the inverse Gaussian density function with mean c and variance 1 when $\theta \to \infty$ and $\lambda \to 1$ such that the product $\theta(1 - \lambda) = c^2$.

9.13 Exercises

9.1 Let X be a discrete random variable which has a generalized Poisson distribution with parameters $(\theta t, \lambda t)$. If $E(X - \mu_1')^k$ is denoted by μ_k, show that

$$\mu_{k+1} = k\theta(1-\lambda)^{-3}\,\mu_{k-1} + (1-\lambda)^{-1}\left\{\frac{d}{dt}\,\mu_k(t)\right\}_{t=1}.$$

By using the above relation, verify the first six central moments given by equations (9.9) and (9.13).

9.2 If the kth negative integer moment of the GPD is denoted by

$$\Phi_k(\theta, r) = E\left[(X + r)^{-k}\right],$$

show that

$$\Phi_2(\theta,\, \theta/\lambda) = \frac{\lambda^2}{\theta^2} - \frac{\lambda^3}{\theta(\theta+\lambda)} - \frac{\lambda^3}{(\theta+\lambda)^2} + \frac{\lambda^4}{(\theta+\lambda)(\theta+2\lambda)}.$$

Find a corresponding expression for the restricted GPD.

9.3 Suppose the probability distribution of finding X bacteria in a given space is denoted by $P_x(\theta, \lambda)$. Suppose further that the mean $\mu(\theta, \lambda)$ of X is increased by changing the parameter λ to $\lambda + \Delta\lambda$ in such a way that

$$\frac{dP_0(\theta, \lambda)}{d\lambda} = 0$$

and

$$\frac{dP_x(\theta, \lambda)}{d\lambda} = -x\, P_x(\theta, \lambda) + \frac{(x-1)\theta}{\theta+\lambda}\, P_{x-1}(\theta + \lambda, \lambda)$$

for all integral values of $x > 0$ with the initial conditions $P_x(\theta, 0) = e^{-\theta}\, \theta^x/x!$ for all values of x. Show that $P_x(\theta, \lambda)$ is a GPD.

9.4 Suppose the initial number k of customers is a Poisson random variable with mean θ per unit service interval and the subsequent arrivals are also Poissonian with mean λ per unit service interval. Prove that the probability distribution of the number of customers served in the first busy period of a single server is the GPD model with parameters (θ, λ).

9.5 Verify the asymptotic biases, variances, and covariance of the moment estimators as given in the results (9.27)–(9.31).

9.6 If a nonnegative GP variate Z is subdivided into two components X and Y in such a way that the conditional distribution $P(X = k, Y = z - k|Z = z)$ is QBD-II with parameters (z, p, θ), show that the random variables X and Y are independent and that they have GP distributions.

9.7 Suppose X follows the restricted GPD with parameters θ and α. Show that a recurrence relation between the noncentral moments is given by

$$\mu'_{k+1} = \mu'_1 \left\{ \mu'_k + \frac{d\mu'_k}{d\theta} \right\}, \quad k = 0, 1, 2, \ldots .$$

Also, obtain a corresponding recurrence relation between the central moments

$$\mu_k, \quad k = 2, 3, 4, \ldots .$$

9.8 Suppose that the probability of buying a product by a person is small and the number of persons is very large. If each buyer of the product becomes its advertiser for a short time in his or her town, which has a large population, show by using the principle of branching process that the total number of persons who will become advertisers will be given by the GPD.

9.9 Draw the graphs of the generalized Poisson distribution for the following sets of parameter values:
(a) $\theta = 8,\ \lambda = -0.1$; (b) $\theta = 8,\ \ \lambda = 0.2$; (c) $\theta = 8,\ \ \lambda = 0.8$;
(d) $\theta = 8,\ \lambda = -2.0$; (e) $\theta = 16,\ \lambda = -2.0$; (f) $\theta = 16,\ \lambda = -3.0$.

9.10 A textile mill produces bolts of cloth of length L. Let X be the actual length of each bolt. If $X \geq 1$, the bolt is sold for $\$A$ and if $X < L$, the bolt is sold as scrap at a price sx, where s is fixed and x is the observed value of X. If the production cost is $c_0 + cx$ dollars, where c_0 and c are the cost constants. Find the expectation $\mathrm{E}(P(X))$ where $P(X)$ is the profit function of the bolt of cloth when
(a) X is a Poisson random variable with mean θ, and
(b) X is a GP random variable with parameters θ and λ.
Find the maximum value of $\mathrm{E}(P(X))$ as θ increases in the two cases. (Hint: See Tripathi, Gupta, and Gupta, 1986.)

9.11 Use the recurrence relation in Exercise 9.7 to determine the mean, the second, third, and fourth central moments of the restricted GPD. Obtain a measure of skewness and a measure of kurtosis. Determine the parameter values for which the restricted GPD is negatively skewed, positively skewed, leptokurtic, and platykurtic.

9.12 Suppose $s = 0$, $\beta = m - 1$, and n are very large, and $b/(b + \omega) = p$ is very small such that $np = \theta$ and $mp = \lambda$ in Prem distribution in (5.14). Show that the Prem distribution approaches the generalized Poisson distribution in (9.1).