

ESTIMASI BENTUK FUNGSIONAL KURVA LORENZ BERDASARKAN DATA PENGELUARAN RUMAH TANGGA (studi kasus Provinsi Papua)

Oleh: Muhammad Fajar

Kasie Statistik Sosial BPS Kab. Waropen

Abstraksi

Tujuan dari penelitian ini adalah untuk mengetahui distribusi pengeluaran rumah tangga di Provinsi Papua dan mengungkapkan fungsi kurva Lorenz berdasarkan distribusi pengeluaran tersebut. Data yang digunakan dalam studi ini adalah data susenas yang dianalisis secara deskriptif dan kuantitatif. Alat analisis yang digunakan adalah probability-probability plot dan rumus pembentuk kurva Lorenz. Kesimpulan dari paper ini adalah bahwa distribusi pengeluaran rumah tangga di Provinsi Papua mengikuti distribusi log normal dan fungsi kurva Lorenz-nya:

$$L(z) = \Phi(\Phi^{-1}(z) - 0.69092)$$

BAB I PENDAHULUAN

1.1 Latar Belakang

Kesejahteraan ekonomi suatu masyarakat pada suatu wilayah salah satunya tergantung dari pendapatan masyarakat. Dengan berbedanya pekerjaan pada setiap orang membuat pendapatan yang diterima berbeda pula sehingga mengakibatkan tingkat heterogenitas ekonomi masyarakat. Hal tersebut membuat Pemerintah harus berhati-hati dalam membuat suatu kebijakan di bidang ekonomi. Oleh karena itu, kondisi distribusi pendapatan atau pengeluaran konsumsi harus diketahui sebagai dasar perencanaan sehingga tepat dalam mengambil kebijakan.

Salah satu alat untuk melihat distribusi pendapatan atau pengeluaran konsumsi adalah kurva Lorenz. Kurva Lorenz adalah salah satu alat analisis untuk mengetahui terdistribusinya pendapatan, kekayaan, ataupun pengeluaran masyarakat. Kurva Lorenz didefinisikan hubungan antara kumulatif proporsi pendapatan dengan kumulasi proporsi pendapatan yang diterima individu. Kurva Lorenz adalah kurva dibawah garis ideal yang menghubungkan titik (0,0) dan titik (1,1). Telah banyak dilakukan penelitian mengenai model dari kurva Lorenz baik dari data kelompok maupun data tunggal seperti yang dilakukan oleh Kakwani and Podder (1973, 76), Rasche et all (1980), Kakwani (1980), Gupta (1984), Ortega (1991), dan Chotikapanich (1993). Tersedianya banyak kemungkinan bentuk persamaan kurva Lorenz memunculkan pertanyaan bentuk mana yang dapat diterapkan pada data pendapatan atau pengeluaran. Salah satu jalan untuk menentukan model yang tepat antara lain dengan melakukan fitting data pada suatu distribusi statistic tertentu. Dengan mengetahui bentuk fungsional dari kurva Lorenz, kita dapat menurunkan berbagai rumus dari ukuran ketidakmerataan, salah satunya indeks gini.

Oleh karena itu dalam paper ini akan membahas bagaimana mengestimasi bentuk fungsional dari kurva Lorenz dengan cara tidak langsung yakni dari kumulatif proporsi dari pendapatan dengan golongan masyarakat yang menerima pendapatan tetapi lebih pada

akarnya dahulu yaitu dengan menyelidiki distribusi alamiah dari data pengeluaran rumah tangga barulah dibentuk kurva Lorenz. Dipilihnya provinsi papua sebagai studi kasus dikarenakan penelitian pertama tentang kurva Lorenz dan mengoptimalkan pemanfaatan data modul Konsumsi Susenas.

1.2 Permasalahan

Berdasarkan uraian sebelumnya, timbul beberapa pertanyaan, antara lain sebagai berikut:

1. Apa distribusi statistik (*probability density function*) dari data pengeluaran rumah tangga?
2. Apa bentuk fungsional kurva Lorenz yang terjadi?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Untuk mengetahui perkembangan pengeluaran rumah tangga di Provinsi Papua
2. Untuk mengetahui distribusi statistik (*probability density function*) dari data pengeluaran rumah tangga.
3. Untuk mengetahui bentuk fungsional dari kurva Lorenz yang terjadi

BAB II LANDASAN TEORI

Distribusi Probabilitas Data Kontinu

Misalkan X adalah variabel acak kontinyu. Maka distribusi probabilitas atau fungsi densitas peluang dari X adalah sebuah fungsi $f(x)$. Jika ada dua bilangan yakni a dan b dimana $a \leq b$, maka nilai peluang (probabilitas) pada interval $[a,b]$ adalah:

$$P(a \leq X \leq b) = \int_a^b f(x) dx \quad \dots (1)$$

Fungsi Distribusi Kumulatif

Cumulative distribution function $F(x)$ atau fungsi distribusi kumulatif untuk variabel acak kontinyu X didefinisikan untuk setiap nilai x adalah sebagai berikut:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y) dy \quad \dots (2)$$

Jenis - jenis distribusi Variabel Acak Kontinyu

Distribusi Normal

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \quad -\infty < x < \infty$$

Jika distribusi normal adalah $N(\mu = 0, \sigma = 1)$ sehingga:

$$f(x; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Distribusi Gamma

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$\alpha > 0, \beta > 0$

Distribusi Eksponential

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\lambda > 0$$

Distribusi *Chi Squared*

$$f(x; \nu) = \begin{cases} \frac{1}{2^{\nu/2} \Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$\nu > 0, \nu \text{ adalah integer}$

Distribusi Weibull

$$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$\alpha > 0, \beta > 0$

Distribusi Log Normal

$$f(x; \mu, \sigma) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma x} e^{-[\ln(x) - \mu]^2 / (2\sigma^2)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$
$$\sigma > 0, \mu > 0$$

Distribusi Half Normal

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-[|x| - \mu]^2 / (2\sigma^2)}, -\infty < x < \infty$$
$$\sigma > 0, \mu > 0$$

Distribusi Beta

$$f(x; \alpha, \beta, A, B) = \begin{cases} \frac{1}{B - A} \cdot \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \left(\frac{x - A}{B - A}\right)^{\alpha-1} \left(\frac{B - x}{B - A}\right)^{\beta-1} & A \leq x \leq B \\ 0 & \text{otherwise} \end{cases}$$

Distribusi Uniform

$$f(x; A, B) = \begin{cases} \frac{1}{B - A} & A \leq x \leq B \\ 0 & \text{otherwise} \end{cases}$$

Distribusi Rayleigh

$$f(x; \theta) = \frac{x}{\theta} e^{-x^2/(2\theta)} \quad x > 0$$

Dimana: θ adalah scale parameter, $\theta > 0$

Distribusi Pareto

$$f(x; k, \theta) = \begin{cases} \frac{k \cdot \theta^k}{x^{k+1}} & x \geq \theta \\ 0 & x < \theta \end{cases}$$

Distribusi Laplace

$$f(x; a, b) = \frac{1}{2b} e^{-|x-a|/b}, \quad -\infty < x < \infty$$

a adalah *location parameter*

b adalah *scale parameter* diestimasi dari data

Distribusi Logistic

$$f(x; a, b) = \frac{1}{b} e^{-(x-a)/b} (1 + e^{-(x-a)/b})^{-2}, \quad b > 0, -\infty < x < \infty, -\infty \leq a \leq \infty$$

a adalah *location parameter*

b adalah *scale parameter* diestimasi dari data

Distribusi Extreme Value tipe I

$$f(x; a, b) = \frac{1}{b} e^{[-(x-a)/b]} e^{-e^{[-(x-a)/b]}}, \quad b > 0, -\infty < x < \infty$$

a adalah *threshold parameter*

b adalah *scale parameter* diestimasi dari data

Distribusi Student t

Jika Z adalah random variabel berdistribusi Normal (0,1) dan V adalah random variabel $\chi^2(r)$
Maka:

$$t = \frac{Z}{\sqrt{V/r}}$$

Dimana: $r > 0$ dan r adalah *integer*.

Karakteristik Kurva Lorenz

Jika pendapatan atau pengeluaran X adalah suatu variabel acak dari suatu fungsi densitas probabilitas yakni $f(x)$, maka *cumulative distribution function* dari $f(x)$ didefinisikan:

$$F(x) = P(X \leq x) = \int_0^x f(x) dx \quad \dots (3)$$

Dimana, $F(x)$ dapat diinterpretasikan sebagai proporsi dari unit yang mempunyai pendapatan/pengeluaran kurang dari atau sama dengan x . nilai dari $F(x)$ bervariasi dari 0 hingga 1. Selanjutnya jika diasumsikan bahwa mean dari pendapatan adalah μ dari $f(x)$. Maka:

$$\mu = \int f(x) dx \quad \dots (4)$$

Lalu momen pertama dari fungsi distribusi X adalah

$$F_1(x) = \frac{1}{\mu} \int_0^x Xf(x) dx \quad \dots (5)$$

Dimana, nilai dari $F_1(x)$ bervariasi dari 0 hingga 1. $F_1(x)$ dapat diinterpretasikan sebagai andil proporsional total pendapatan/pengeluaran dari unit yang mempunyai pendapatan/pengeluaran kurang dari atau sama dengan x . Maka kurva Lorenz diartikan sebagai hubungan antara variabel $F(x)$ dengan $F_1(x)$. Gastwirth (1971) mendefinisikan kurva Lorenz sebagai berikut:

$$L(z) = \frac{1}{\mu} \int_0^z F^{-1}(x) dx \quad \dots (6)$$

$L(z)$ adalah share pendapatan/pengeluaran dari kelas ke- i dan z adalah proporsi unit pendapatan/pengeluaran dari kelas ke- i . Sebagai contoh, bila 30 persen penduduk memiliki 10 persen dari total pendapatan, maka $L(z) = 0.10$ dan $z = 0.30$.

Kakwani (1980) memberikan catatan tentang karakteristik kurva Lorenz ($L(z)$), sebagai berikut:

1. $L(z) = 0$ jika $z = 0$
2. $L(z) = 1$ jika $z = 1$
3. Turunan pertama dari $L(z) \geq 0$, untuk $0 \leq z \leq 1$ atau $L'(z) \geq 0$ untuk $0 \leq z \leq 1$
4. Turunan kedua dari $L(z) \geq 0$, untuk $0 \leq z \leq 1$ atau $L''(z) \geq 0$ untuk $0 \leq z \leq 1$
5. $L(z) \leq z$ untuk $0 \leq z \leq 1$

BAB III METODOLOGI

Data yang digunakan dalam studi ini adalah data susenas modul konsumsi tahun 2013 Provinsi Papua.

- Probability – Probability Plot

Probability – Probability Plot adalah alat satu alat statistic berupa grafik antara yang menggambarkan titik-titik dari kumulatif probabilitas empiris data terhadap kumulatif probabilitas suatu distribusi teoretis X. Sehingga apabila secara kecenderungan titik-titik kumulatif probabilitas dari data empiris sesuai dengan kumulatif probabilitas distribusi X, maka dapat dikatakan bahwa data empiris mengikuti distribusi X. Pada grafik terdapat garis lurus (garis yang terbentuk dari nilai yang sama antara kumulatif distribusi dari data dengan kumulatif distribusi teoretis X), jika plot – plot jatuh pada garis lurus (mendekati bentuk garis lurus) tersebut maka dapat disimpulkan bahwa data observasi mengikuti suatu distribusi X. Jika plot-plot tidak jatuh pada garis lurus tersebut (tidak mendekati bentuk garis lurus), maka dapat disimpulkan bahwa data observasi tidak mengikuti distribusi X.

- Tahap – Tahap Penelitian

1. Membuat Probability – Probability Plot dari berbagai distribusi kontinu terhadap data pengeluaran rumah tangga.
2. Untuk menentukan distribusi dari pengeluaran rumah tangga adalah dengan cara melihat Pada grafik terdapat garis lurus (garis yang terbentuk dari nilai yang sama antara kumulatif

distribusi dari data dengan kumulatif distribusi teoretis X), jika plot – plot jatuh pada garis lurus (mendekati bentuk garis lurus) tersebut maka dapat disimpulkan bahwa data observasi mengikuti suatu distribusi X.

3. Jika telah diketahui distribusi teoretis dari pengeluaran rumah tangga, maka untuk menentukan persamaan kurva Lorenz adalah:

- a. Menentukan cdf dari fungsi densitas probabilitas $f(x)$ dengan rumus:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y) dy \quad \dots (7)$$

- b. Kemudian tentukan fungsi dari invers $F(x)$ yakni $F^{-1}(x)$

4. Fungsi kurva Lorenz adalah:

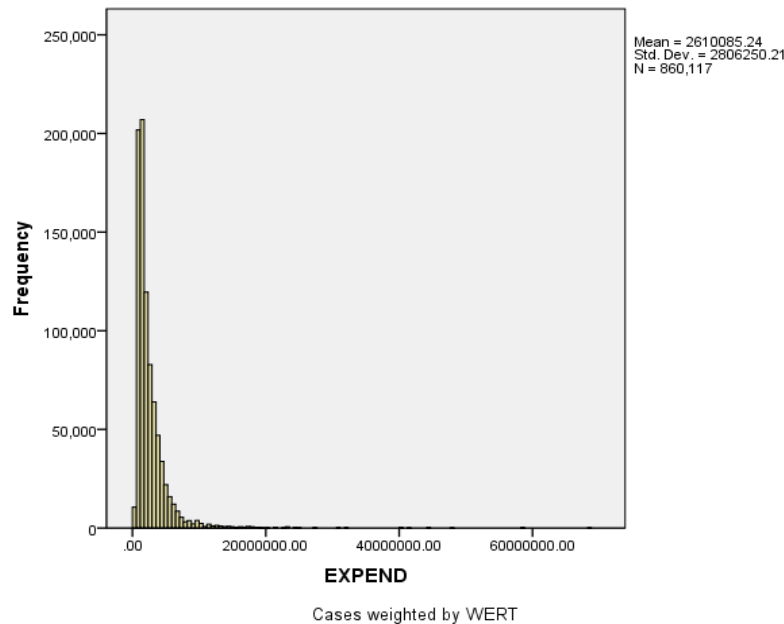
$$L(z) = \frac{1}{\mu} \int_0^z F^{-1}(x) dx \quad \dots (8)$$

BAB IV ANALISIS DAN PEMBAHASAN

Statistik Deskriptif Pengeluaran Rumah Tangga

Rata-rata pengeluaran rumah tangga di Provinsi Papua adalah Rp 2,610,085.2390 dengan standar error \pm Rp 3025.85074. Nilai maksimum pengeluaran rumah tangga di Rp 68,404,945.54 dan minimum pengeluaran rumah tangga Rp 258336.67. Modus pengeluaran rumah tangga di Provinsi Papua sebesar Rp 3,504,095.24. Jika kita lihat histogram yang terbentuk adalah tidak simetri atau *positively skewed*.

Gambar 1. Histogram Pengeluaran Rumah Tangga di Provinsi Papua Tahun 2013



Sumber: Data Susenas, diolah.

Distribusi Pengeluaran Rumah Tangga

Berdasarkan hasil *probability-probability plot*¹ dari berbagai distribusi random variabel kontinu yang dibahas pada bab sebelumnya, ternyata data pengeluaran rumah tangga mengikuti distribusi log normal karena hanya pada plot PP atas distribusi log normal mendekati garis lurus (lihat lampiran). Atas dasar tersebut, menurut kakwani (1980a) kurva Lorenz yang terbentuk dari distribusi log normal adalah simetris.

Hasil estimasi parameter pada distribusi log normal dari pengeluaran rumah tangga adalah:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right]} \quad \dots (9)$$

Dimana:

Estimasi dari $\sigma = 0.69092$ dan $\mu = 14.4973$

Sehingga fungsi log normal di atas dapat ditulis sebagai berikut:

$$f(x) = \frac{1}{1.73188x} e^{\left[-\frac{(\ln x - 14.4973)^2}{0.47737}\right]} \quad \dots (10)$$

Beberapa karakteristik yang menguntungkan dari distribusi log normal, yaitu:

1. Andaikan semua pengeluaran diubah secara proporsional oleh sebuah faktor pengali acak, dimana berbeda disetiap individunya dan mengikuti proses *Gaussian*. Maka distribusi pengeluaran dari populasi akan menuju log normal, bila prosesnya aktif dalam periode yang cukup lama.
2. Fitting log normal cukup baik untuk banyak data.
3. Kurva Lorenz yang terbentuk adalah simetris
4. Ukuran ketidakmerataan (*inequality measures*) tergantung pada parameter σ .

Fungsi Kurva Lorenz

Menurut Lubrano (2014) fungsi kurva Lorenz yang terbentuk dari distribusi log normal adalah sebagai berikut (sesuai dengan rumus pada persamaan (8)):

$$L(z) = \Phi(\Phi^{-1}(z) - \sigma) \quad \dots (11)$$

Dimana:

Φ adalah cdf (*cumulative distribution function*) normal baku:

$$\Phi(z) = Prob(Z < z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad \dots (12)$$

$\Phi(z) = Prob(Z < z)$

$\Phi^{-1}(z)$ = invers dari $\Phi(z)$

z adalah proporsi kelompok atau individu yang memiliki pengeluaran tertentu.

Sehingga untuk kasus pengeluaran rumah tangga di Provinsi Papua pada tahun 2013, maka fungsi kurva Lorenz yang terbentuk adalah:

$$L(z) = \Phi(\Phi^{-1}(z) - 0.69092)$$

¹ Karena probability-probability plot untuk distribusi chi squared tidak terbentuk, maka digunakan QQ plot

KESIMPULAN

Berdasarkan pada pembahasan sebelumnya dapat kita tarik kesimpulan bahwa:

1. Distribusi konsumsi rumah tangga di Provinsi Papua adalah distribusi lognormal, yaitu:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right]}$$

Dimana:

Estimasi dari $\sigma = 0.69092$ dan $\mu = 14.4973$ sehingga menjadi:

$$f(x) = \frac{1}{1.73188x} e^{\left[-\frac{(\ln x - 14.4973)^2}{0.47737}\right]}$$

2. Atas dasar poin (1) maka kurva Lorenz berasal dari fungsi lognormal sehingga:

$$L(z) = \Phi(\Phi^{-1}(z) - 0.69092)$$

REFERENSI

Battistin, E. Blundell, R. and A. Lewbel, 2007, "*Why is Consumption More Log Normal Than Income? Gibrat's Law Revisited*", Working Paper 08/07, Institute for Fiscal Studies, London, and Working Paper 671, Department of Economics, Boston College.

Chotikapanich, D, 1993, "*A Comparison of Alternative Functional Forms for the Lorenz Curve*" *Economics Letters* 41, pp. 21-29.

Chotikapanich, D, Griffiths, B, and Rao, D.S, 2007, "*Estimating and Combining National Income Distributions Using Limited Data*" *Journal of Business and Economic Statistics* 25, pp. 97-109.

Gastwirth, J and Glauber, M, 1976, "*The Interpolation of the Lorenz Curve and Gini Index from Grouped Data*" *Econometrica* 44, pp. 479-483.

Gupta, M.R, 1984, "*Functional Form for Estimating the Lorenz Curve*" *Econometrica* 52, pp. 1313-1314.

Chakravarty, S.R., 1988, "*Extended Gini Indices of Inequality*", *International Economic Review*, 29, 147-156.

Gastwirth, J. L., 1971, "*A General Definition of the Lorenz Curve*", *Econometrica*, 39, 1037-1039.

Gupta, M.R., 1984, "*Functional Form for Estimating the Lorenz Curve*", *Econometrica*, 52, 1313-1314.

Kakwani N., 1980a, "*Income Inequality and Poverty*", *Oxford*: Oxford University Press.

Kakwani N., 1980b, "*On a Class of Poverty Measures*", *Econometrica*, 48, 437-446.

Kakwani N., and Podder, N., (1973), "*On the Estimation of Lorenz Curves from Grouped Observations*", *International Economic Review*, 14-2, 278-29.

Kakwani N., and Podder, N., 1976, "*Efficient Estimation of Lorenz Curve and Associated Inequality Measures from Grouped Observations*", *Econometrica*, 14-2, 278-291.

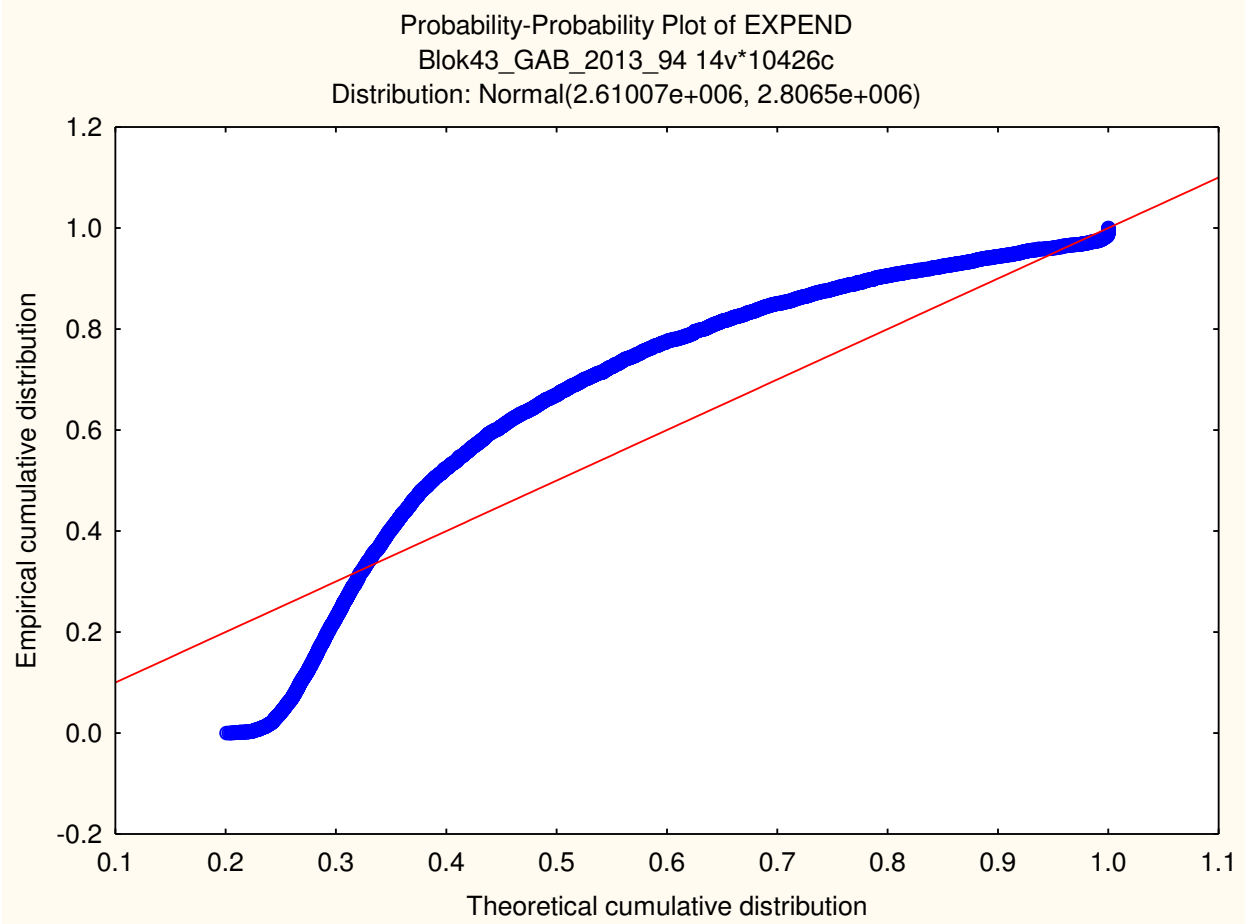
Lorenz, M. O., 1905, "*Methods of Measuring the Concentration of Wealth*", *Journal of American Statistical Association*, 9, 209-219.

Lubrano, Michel, 2014, "*Lecture 4: Lorenz curves, the Gini coefficient and parametric distribution*",

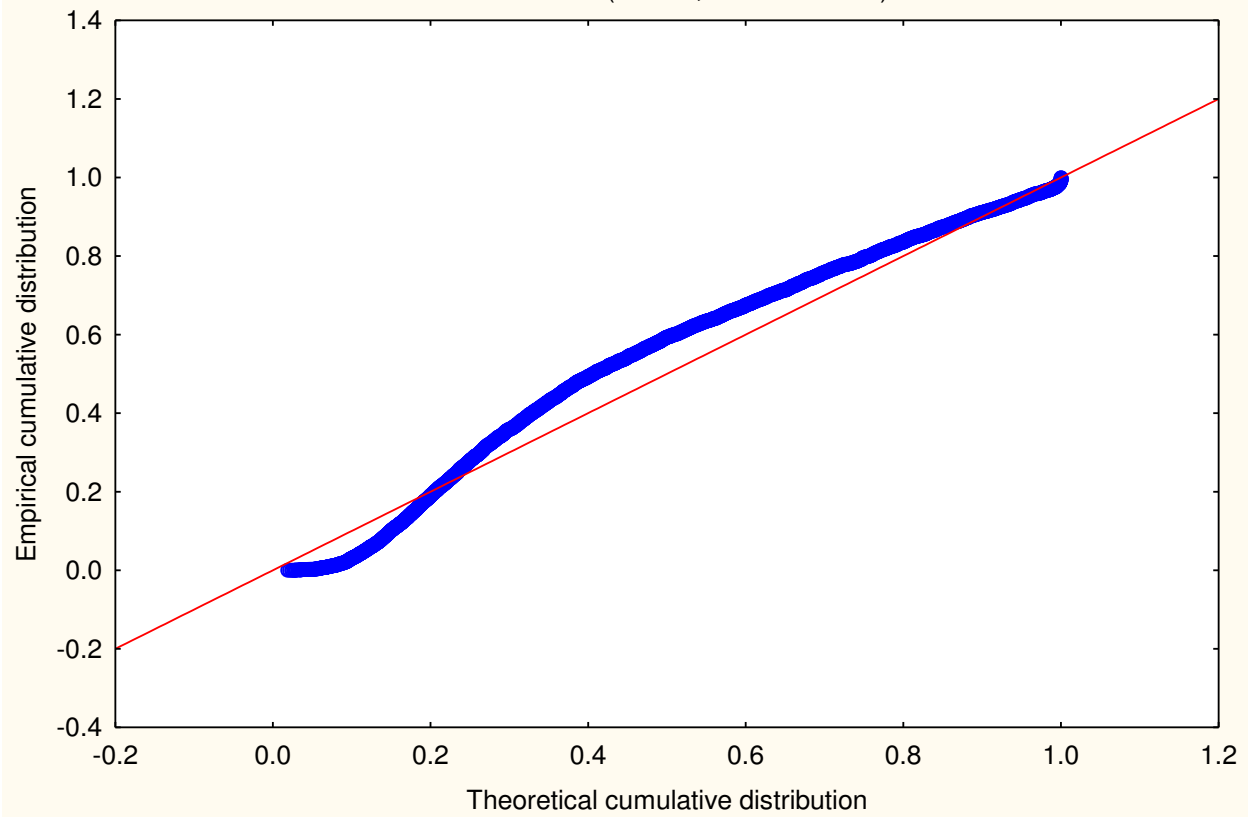
Rasche, R. H., Gaffney, J. Koo, A. Y., and Obst, N., 1980, "*Functional Forms for Estimating the Lorenz Curve*", *Econometrica*, 48, 1061-1062.

LAMPIRAN

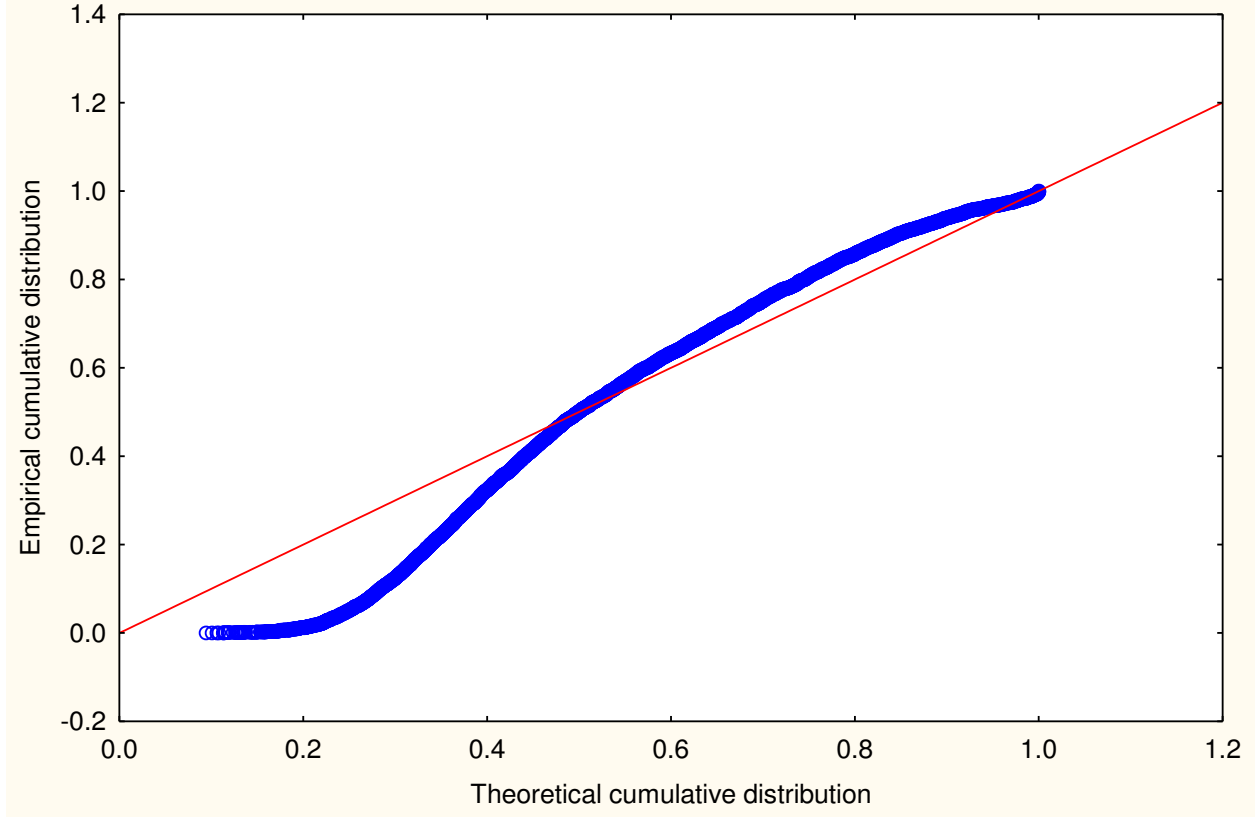
Probability-Probability Plot Pengeluaran Rumah Tangga Provinsi Papua Tahun 2013

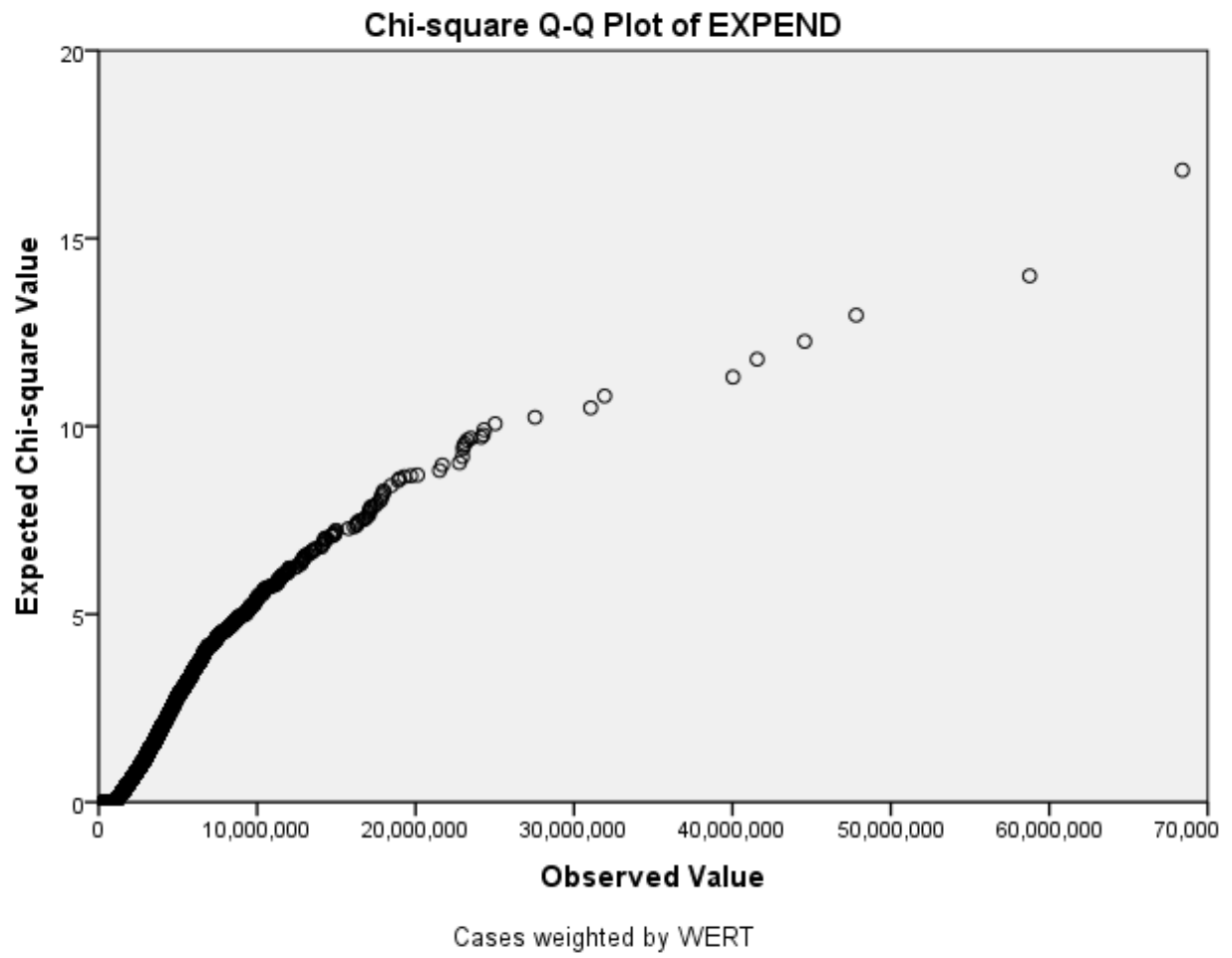


Probability-Probability Plot of EXPEND
Blok43_GAB_2013_94 14v*10426c
Distribution: Gamma(1.9515, 1.33747e+006)

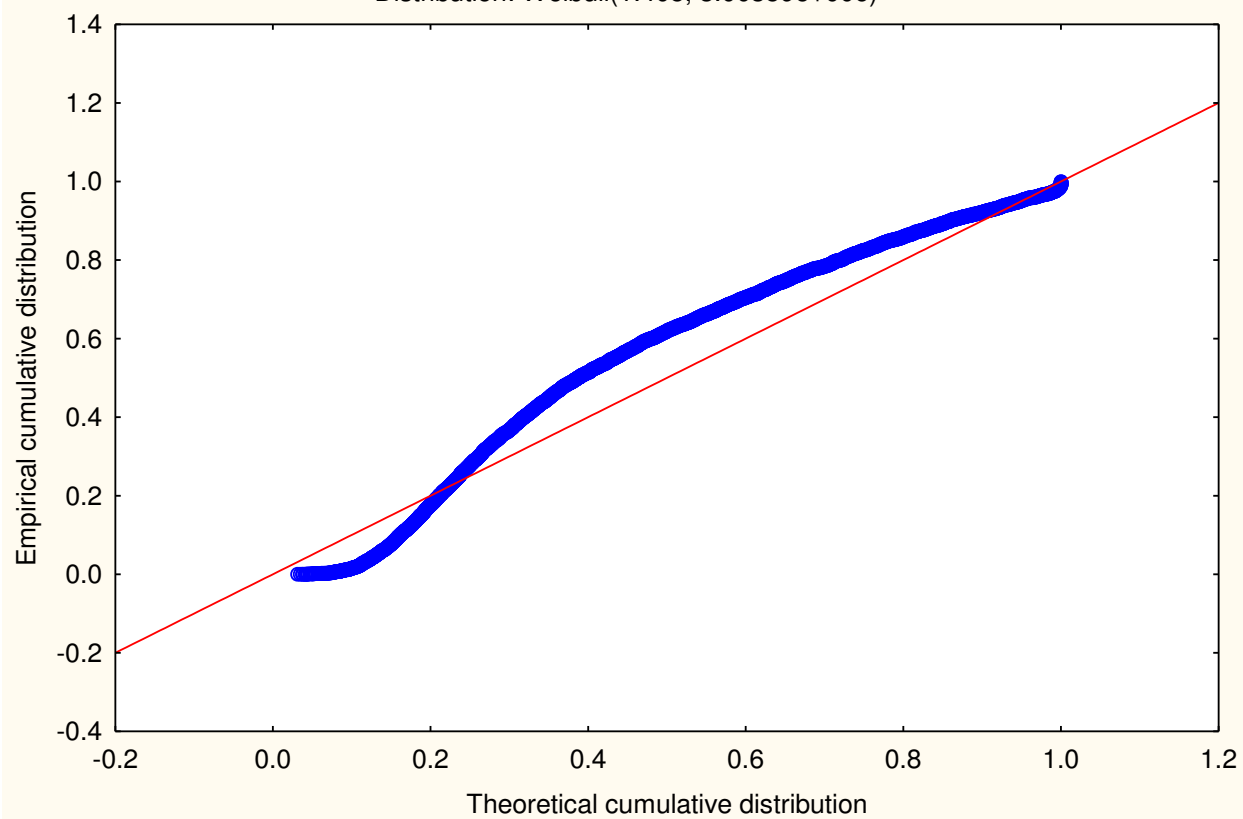


Probability-Probability Plot of EXPEND
Blok43_GAB_2013_94 14v*10426c
Distribution: Exponential(2.61007e+006)

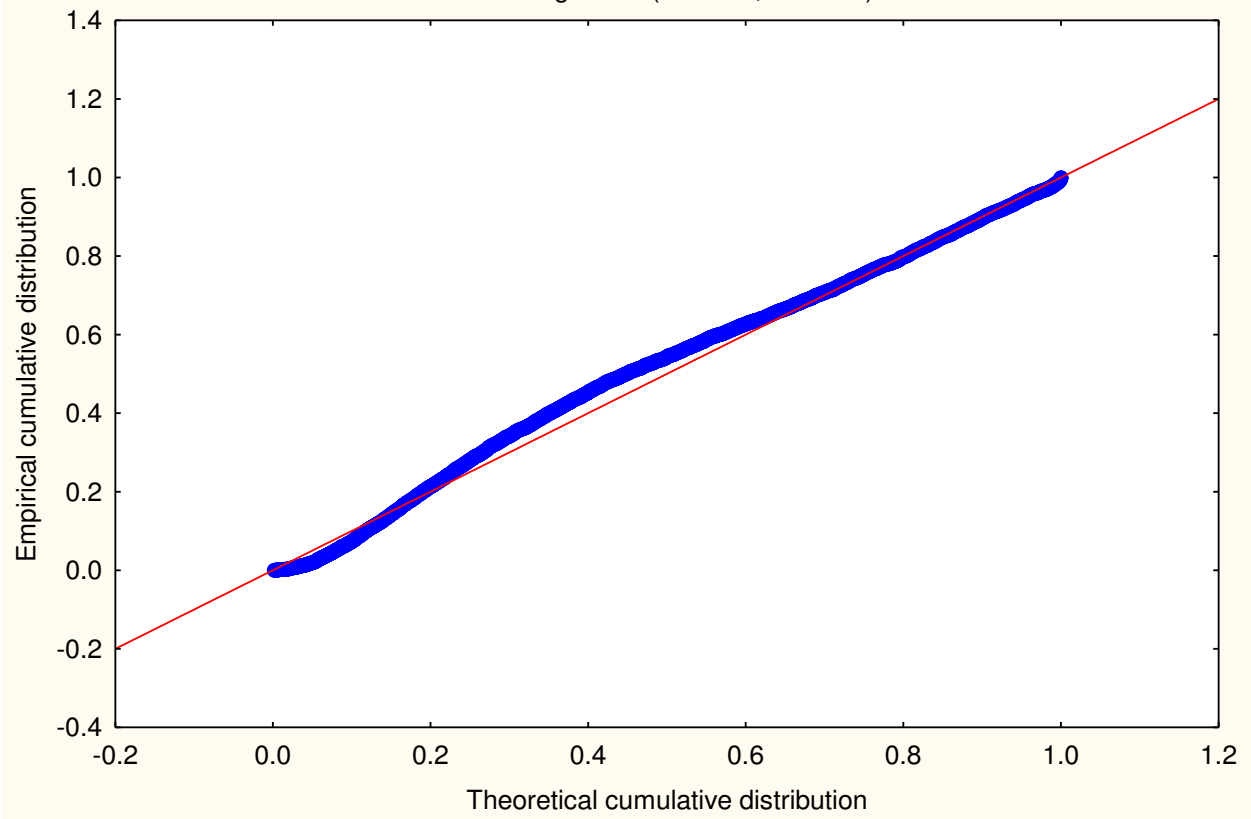




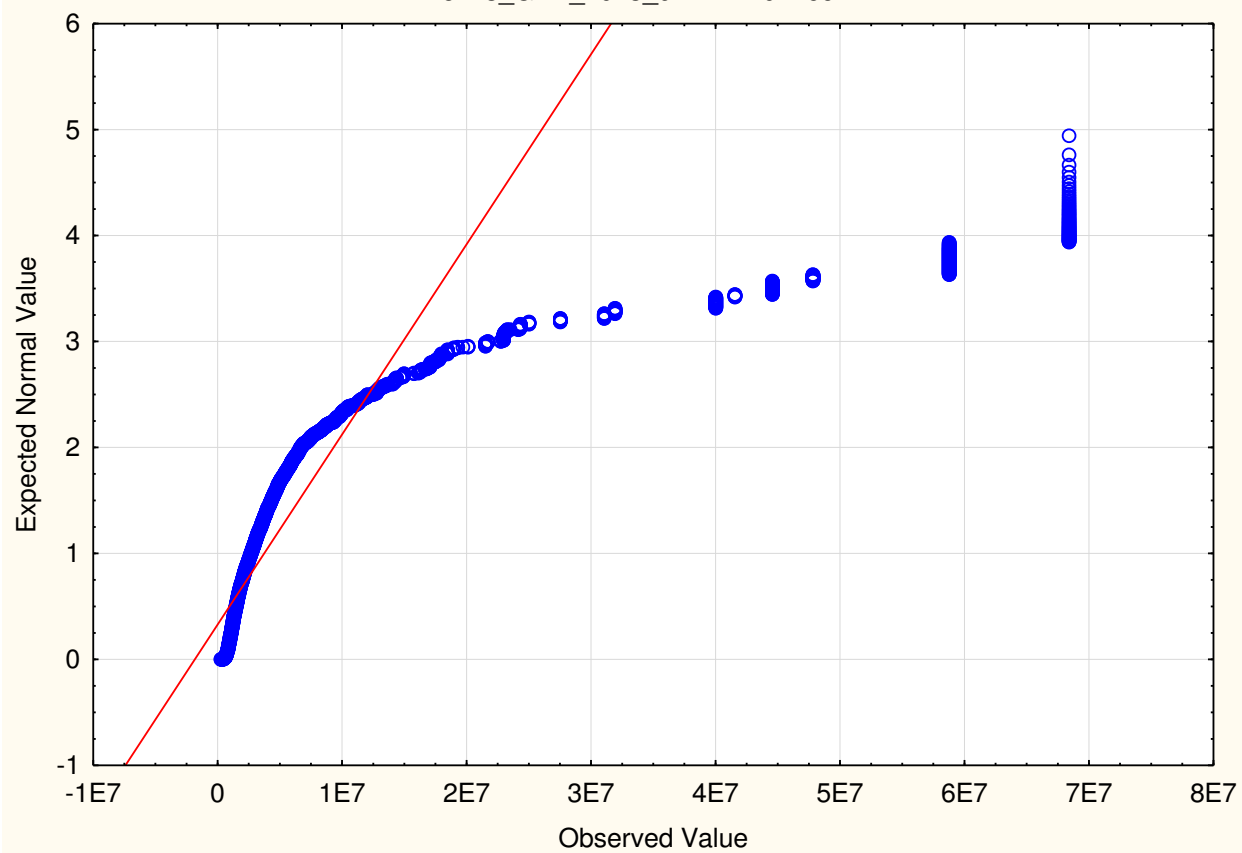
Probability-Probability Plot of EXPEND
Blok43_GAB_2013_94 14v*10426c
Distribution: Weibull(1.405, 3.00359e+006)



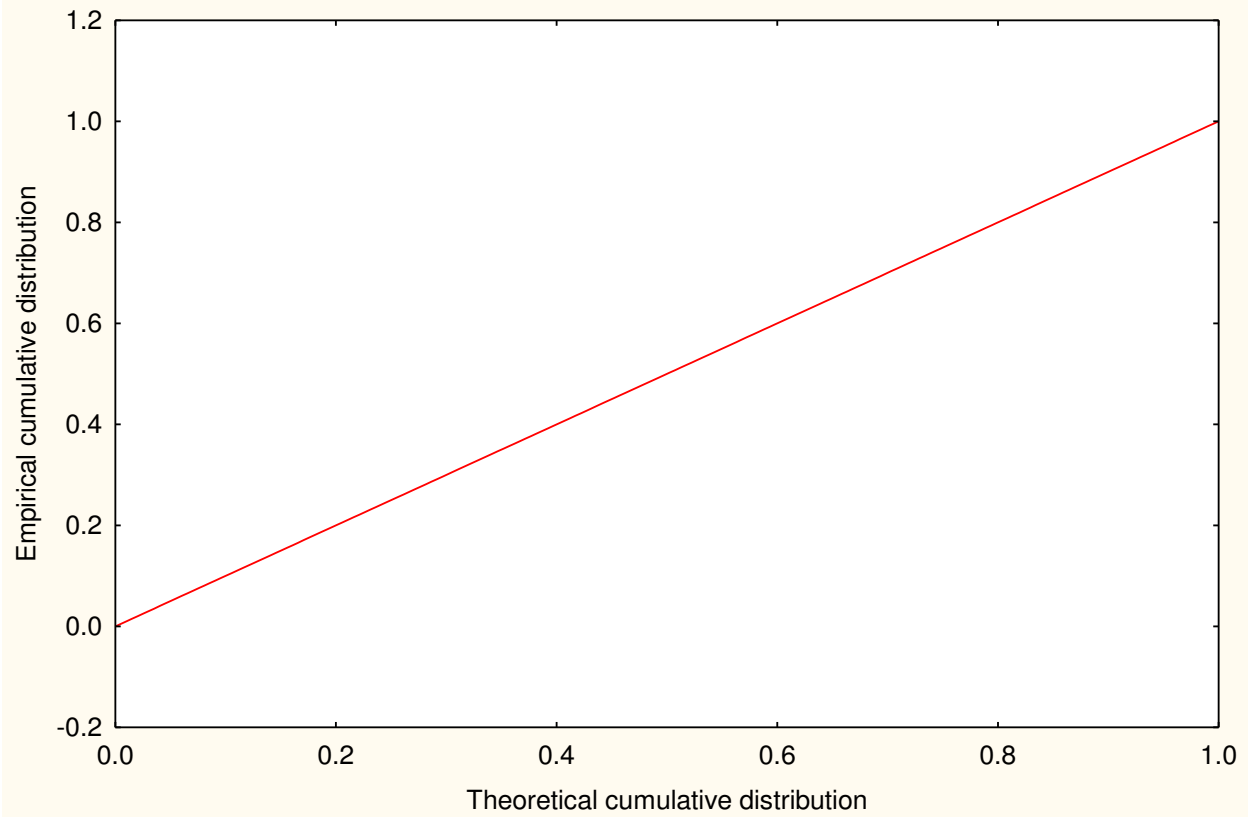
Probability-Probability Plot of EXPEND
Blok43_GAB_2013_94 14v*10426c
Distribution: Lognormal(0.69092, 14.4973)

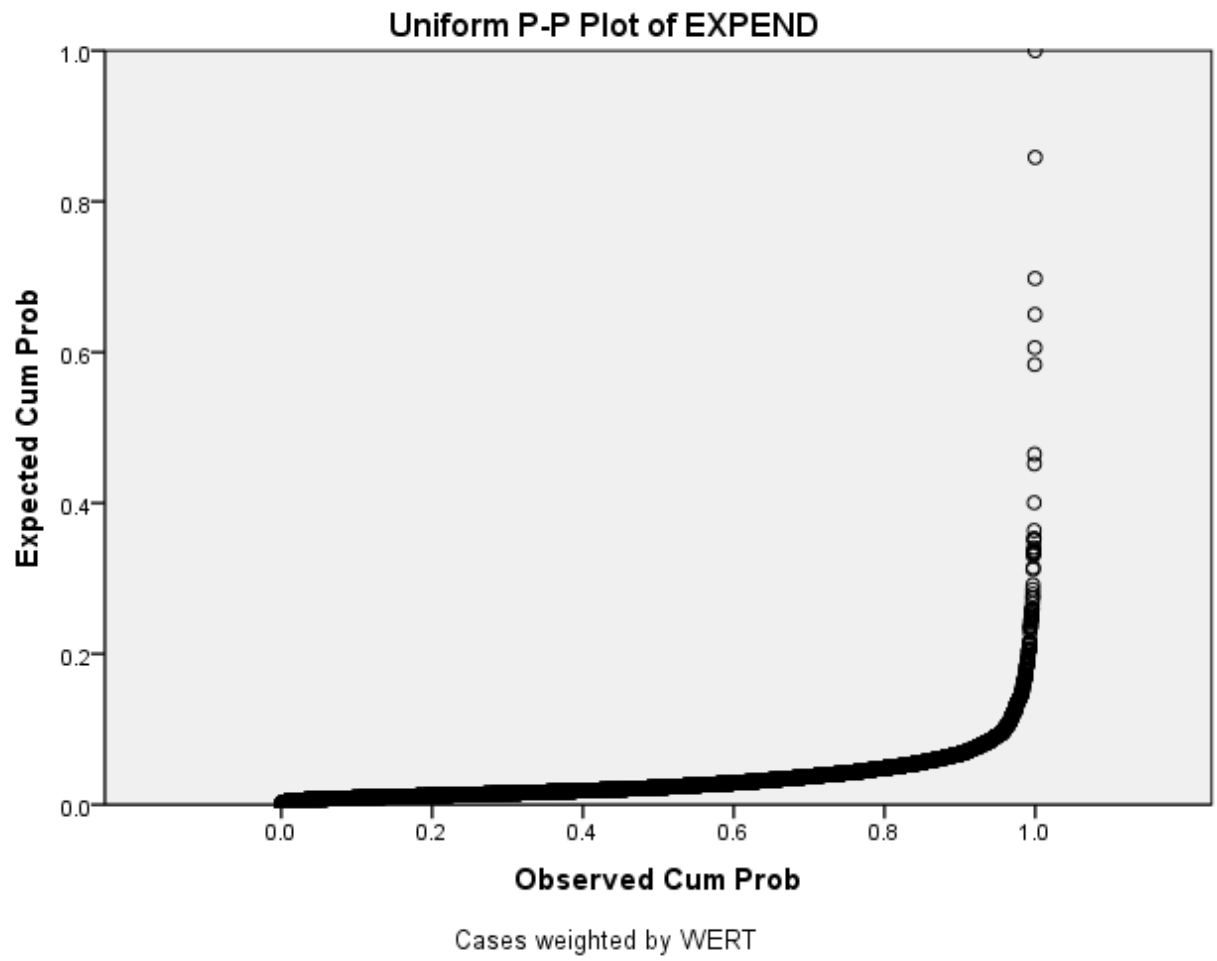


Half-Normal Probability Plot of EXPEND
Blok43_GAB_2013_94 14v*10426c

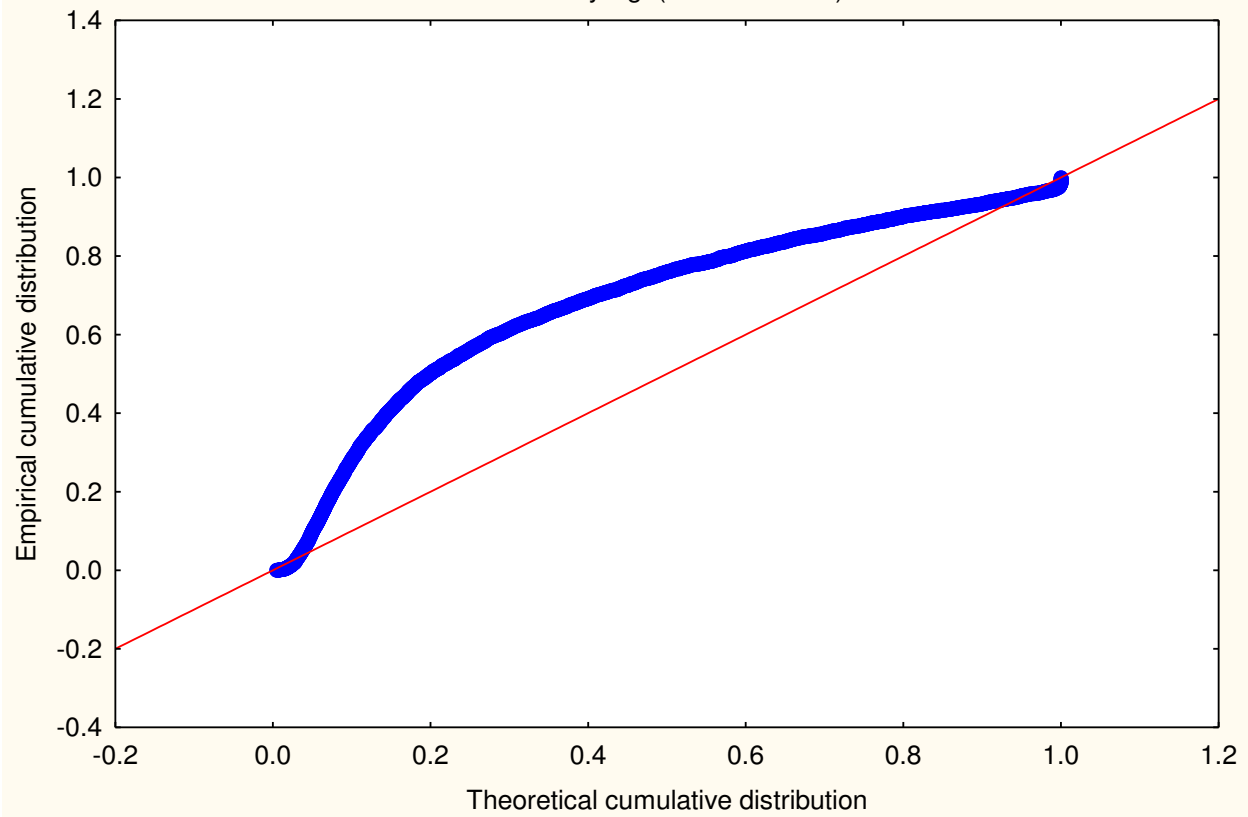


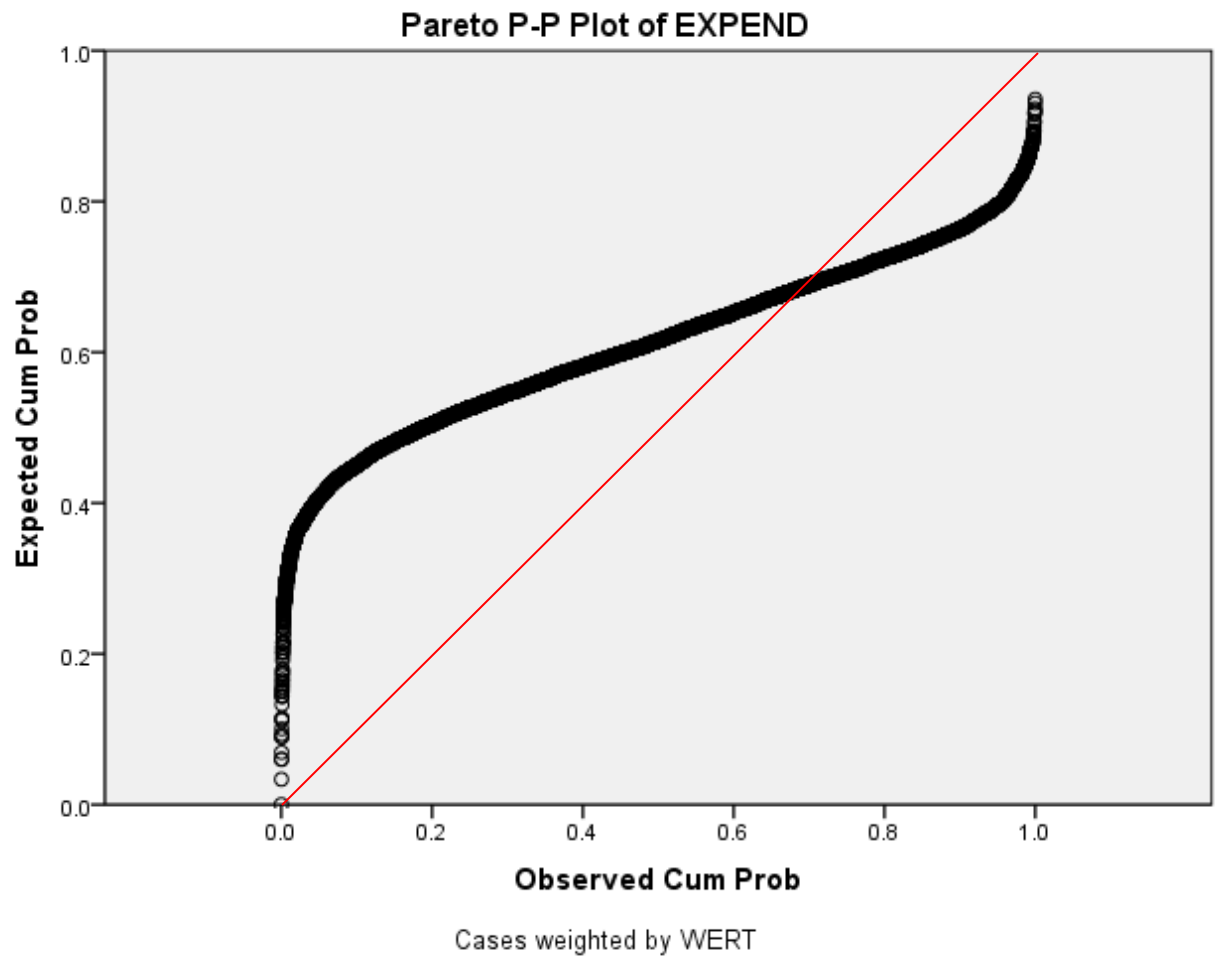
Probability-Probability Plot of EXPEND
Blok43_GAB_2013_94 14v*10426c
Distribution: Beta

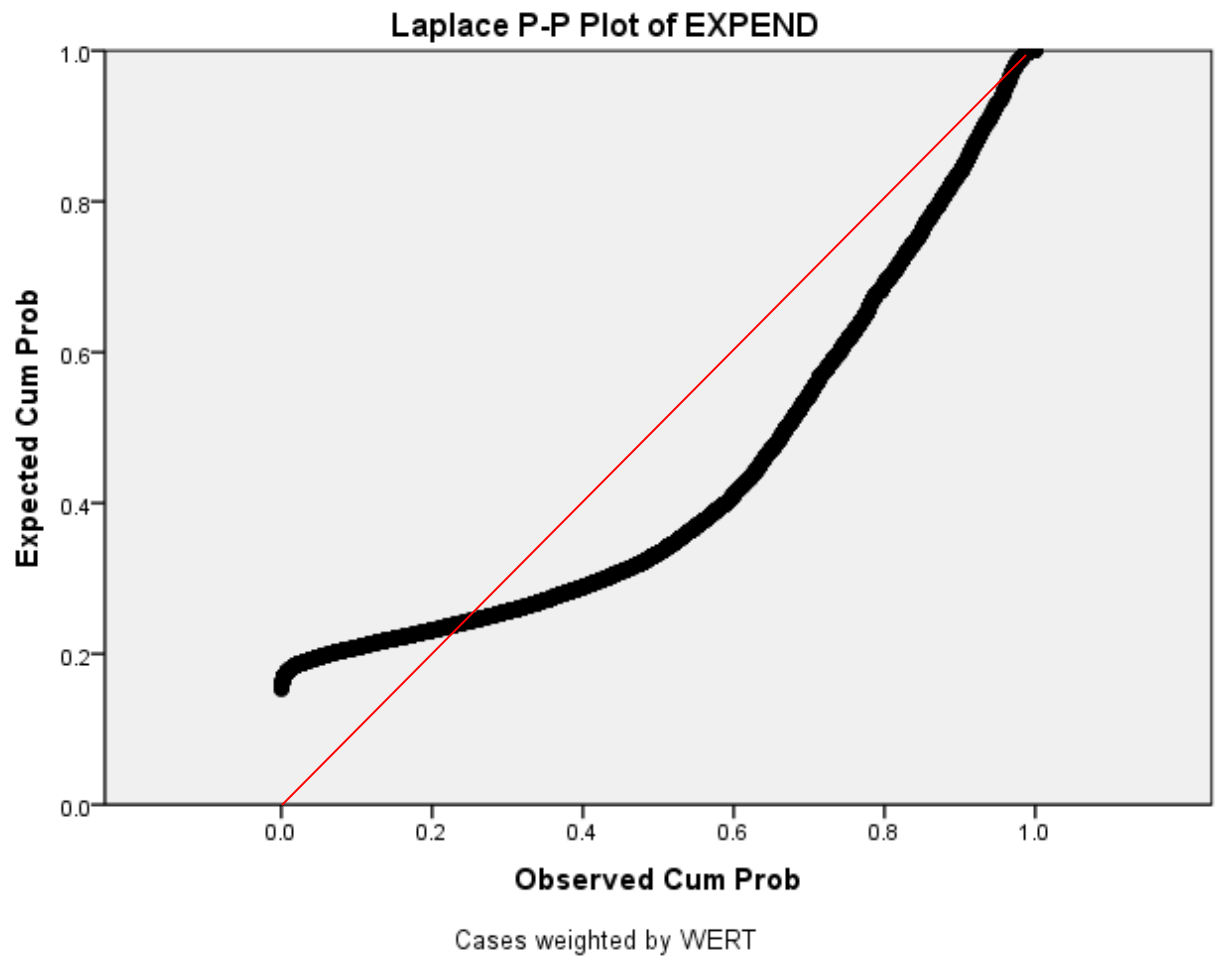


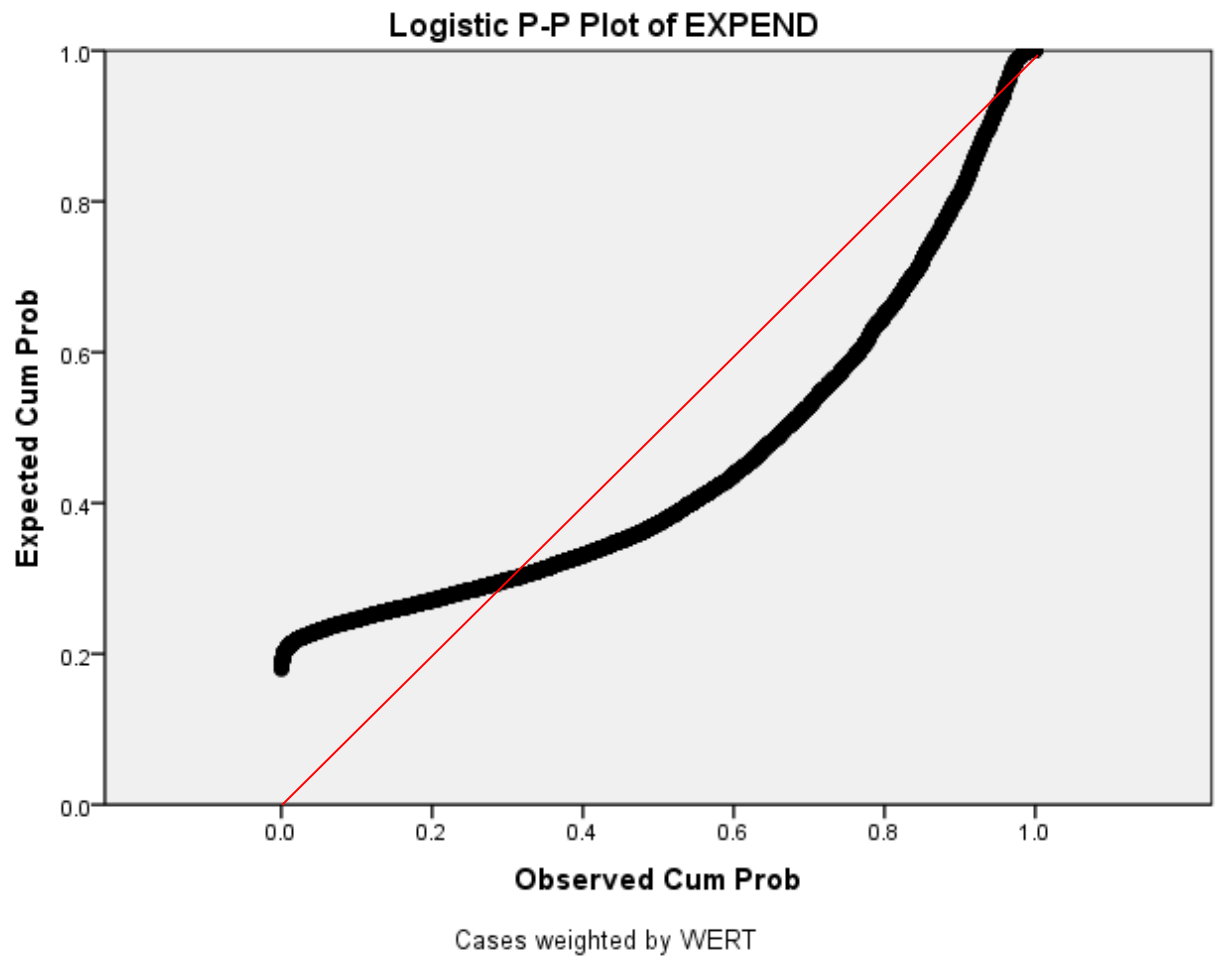


Probability-Probability Plot of EXPEND
Blok43_GAB_2013_94 14v*10426c
Distribution: Rayleigh(2.71006e+006)









Probability-Probability Plot of EXPEND
Blok43_GAB_2013_94 14v*10426c
Distribution: Extreme(1.74514e+006, 1.23959e+006)

