ISBN: 978-979-792-552-9

Peluang Pengembangan Sebaran *T*-*X*

Herlina Hanum

Jurusan Matematika, FMIPA Universitas Sriwijaya.
Program Doktor Statistika Institut Pertanian Bogor

linhanum@gmail.com

A H Wigena, A Djuraidah

Departemen Statistika, FMIPA, Institut Pertanian Bogor

I W Mangku

Departemen Matematika, FMIPA Institut Pertanian Bogor

Abstrak

Perkembangan teori sebaran akhir-akhir ini telah banyak menghasilkan sebaran peluang yang baru. Keluarga sebaran *transformed-transformator* yang populer dengan istilah *T*-*X* merupakan hasil perkembangan terbaru. Sebaran *T*-*X* dikembangkan berbasis fungsi komposisi antara peubah acak *T* dan *X*, misalnya sebaran Beta-Normal, Kumaraswamy *generalized gamma*, dan Gamma-Pareto. Dalam tulisan ini diuraikan aturan pembentukan sebaran *T*-*X*, dan pembentukan sebaran Gamma-Pareto sebagai dasar pengetahuan untuk menciptakan sebaran-sebaran peluang yang baru. Selanjutnya dibahas peluang penelitian yang dapat dilakukan dalam perkembangan sebaran *T*-*X*. Tulisan ini membuka wawasan tentang perkembangan terbaru teori sebaran terutama sebaran *T*-*X* dan materi penelitian yang dapat dilakukan untuk turut serta dalam pengembangan sebaran *T*-*X*.

Kata kunci: pemodelan, peubah acak, Gamma-Pareto, sebaran *T*-*X*

1 Pendahuluan

Teori peluang sebagai dasar dalam statistika inferensia selalu mengalami perkembangan. Perkembangan tersebut dapat dilihat salah satunya dari semakin banyaknya bentuk sebaran peluang. Dalam buku teori peluang tahun 1980-an bentuk sebaran yang ditemui mungkin hanya terbatas pada beberapa sebaran baku. Misalnya sebaran normal, t, gamma, dan F, untuk peubah acak kontinu. Sementara untuk peubah acak diskret ada binomial dan poisson. Saat ini sudah ada gabungan dari peubah acak diskret dan kontinu. Misalnya sebaran beta binomial yang didapat dari aturan bayes. Bentuk sebaran baru yang lain misalnya bentuk-bentuk baru sebaran normal seperti *skew-normal*, *folded normal*, dan *half normal*.

Perkembangan bentuk sebaran tidak terlepas dari berkembangnya tehnik penciptaan sebaran baru. Teknik-teknik yang digunakan saat ini banyak memanfaatkan sebaran yang sudah ada. Teknik yang terbaru adalah pembentukan keluarga sebaran *transformed-transformer* atau disebut sebaran *T-X*. Istilah sebaran *T-X* dipopulerkan oleh Alzaatreh *et al*. (2013). Keluarga sebaran *T-X* menggabungkan dua sebaran peubah acak. Peubah acak pertama, dilambangkan *T*, ditransformasi oleh peubah bebas kedua yang dilambangkan *X*. Istilah lain yang sering digunakan untuk sebaran ini adalah *T-generated*. Artinya keluarga sebaran yang dibangkitkan oleh sebaran *T*. Sebagai contoh *beta-generated* (Eugene *et al*. 2002) yaitu keluarga sebaran yang dibangkitkan dari sebaran beta, dan *gamma-generated* yang dibangkitkan dari sebaran gamma (Alzaatreh *et al*. 2013)

Teknik yang digunakan dalam menciptakan keluarga sebaran *T-X* adalah pembentukan fungsi sebaran kumulatif peubah acak *T-generated* menggunakan fungsi sebaran kumulatif peubah acak *T* pada $t = F(x)$. $F(x)$ adalah fungsi sebaran kumulatif sebarang peubah acak *X*. Untuk keluarga sebaran *beta-generated* fungsi sebaran kumulatifnya adalah

$$G(x) = \int_0^{F(x)} b(t)\, dt$$

dengan $b(t)$ adalah fungsi kepekatan peluang (fkp) sebaran $beta(\alpha, \beta)$, sedangkan fkp -nya

$$g(x) = \frac{1}{B(\alpha,\beta)} f(x) F^{\alpha-1}(x)(1 - F(x))^{\beta-1}$$

Eugene *et al* (2002).

Sejak sebaran beta-normal dikenalkan oleh Eugene *et al* (2002), banyak sebaran beta-X lain sudah dibentuk. Sebaran tersebut antara lain beta-weibull (Famoye *et al*, 2005), beta-eksponensial (Nadarajah dan Kotz, 2005), beta-pareto(Akinsete *et al*, 2008), dan beta-gamma (Kong *et al* 2007). Selain keluarga sebaran beta-X, juga telah dikembangkan keluarga sebaran lain seperti gamma-X, kumaraswamy-X, dan Weibull-X. Keluarga sebaran gamma-X dikembangkan oleh Alzaatreh *et al* (2012, 2013a, 2014). Sementara keluarga sebaran kumaraswamy-X antara lain dikembangkan oleh Jones (2009) dan Cordeiro *et al* (2010). Alzaatreh *et al* (2013a) menguraikan bentuk umum keluarga weibull-X, disusul dengan sebaran weibull-pareto (Alzaatreh *et al*, 2013b)

2 Tujuan dan Alur Penulisan

Pengembangan wawasan perkembangan terkini sebaran peluang sangat penting baik bagi pengajar maupun peneliti statistika dan terapannya. Bagi pengajar sangat penting untuk memperkaya materi pembelajaran. Selain itu juga membuka wawasan peserta didik agar mereka dapat memahami bahwa teori sebaran itu suatu ilmu yang dinamis. Hal ini akan memacu baik pengajar maupun peserta didik untuk terus mengikuti

ISBN: 978-979-792-552-9

perkembangan yang terjadi dari waktu ke waktu. Bagi peneliti, pengembangan wawasan tersebut akan banyak membuka peluang-peluang penelitian yang sesuai dengan perkembangan keilmuan. Tulisan ini dimaksudkan untuk memacu semua pihak yang berhubungan dengan sebaran peluang baik pengajar, peserta didik, peneliti, pengguna dan pihak lain untuk terus mengikuti dan turut serta dalam pengembangan sebaran peluang terutama keluarga sebaran *T-X*.

Dalam tulisan ini pertama sekali dibahas tentang aturan pembentukan sebaran *T-X*. Selanjutnya diberikan contoh pembentukan sebaran *T-X* yaitu sebaran Gamma-Pareto. Pada bagian ketiga dikaji peluang-peluang penelitian yang dapat dilakukan berkaitan dengan sebaran *T-X*.

3 Aturan Pembentukan *T-X*

Pembentukan sebaran *T-X* dimulai dengan pembentukan fungsi sebaran peubah acak *T-X* yaitu

$$G(x) = \int_0^{W(F(x))} r(t)\, dt$$

T suatu peubah acak kontinu yang terdefinisi di *[a,b]* dengan $-\infty \leq a < b \leq \infty$. Sementara *r(t)* adalah fkp peubah acak *T dan F(x)* fungsi sebaran peubah acak *X* Peubah acak *X* boleh bertype diskret atau kontinu. *W(F(x))* suatu fungsi dari *F(x)* yang memenuhi syarat $W(F(x)) \epsilon (0, 1)$, terdiferensialkan, monoton tak turun, $W(F(x)) \to 0$ *jika* $x \to -\infty$ dan $W(F(x)) \to 1$ jika $x \to \infty$ (Alzaatreh *et al*. 2013a).

Bentuk fungsi *W(F(x))* tergantung pada daerah definisi *T*. Daerah definisi T dan fungsi *W(F(x))* diuraikan dalam Alzaatreh *et al*. (2013a) adalah

1. $T \ \epsilon (0, 1)$: $F(x)$ atau $F^{\alpha}(x)$.
2. $T \ \epsilon [a, \infty), a \geq 0$: $-\log(1 - F(x))$, $F(x)/(1 - F(x))$, $-\log(1 - F^{\alpha}(x))$ dan $F^{\alpha}(x)/(1 - F^{\alpha}(x))$, dengan $\alpha > 0$
3. $T \ \epsilon (-\infty, \infty)$: $\log[-\log(1 - F(x))]$, $\log[F(x)/(1 - F(x))]$, $\log[-\log(1 - F\alpha(x))]$, and $\log[F^{\alpha}(x)/(1 - F\alpha(x))]$.

Fkp peubah acak *T-X* didapat dengan menurunkan *G(x)* terhadap *X*. Bentuk umum fkp tersebut adalah (Alzaatreh *et al*. 2013a)

$$g(x) = \left\{\frac{d}{dx}W(F(x))\right\} r\{W(F(x))\}$$

4 Pembentukan Sebaran Gamma-Pareto

Sebaran Gamma-Pareto dikembangkan oleh Alzaatreh *et al* (2012). Sebaran ini ternasuk keluarga sebaran *T-X*, khususnya *Gamma-X*. Artinya sebaran Gamma yang ditransformasi dengan sebaran *X*, dalam hal ini sebaran Pareto. Pembentukan sebaran Gamma-Pareto dimulai dengan menentukan fungsi sebaran kumulatif sebaran *T-X* yang berbentuk

$$G(x) = \int_0^{-\log(1-F(x))} r(t) dt \qquad (1)$$

ISBN: 978-979-792-552-9

dengan *F(x)* adalah fungsi sebaran kumulatif peubah acak *X*, *r(t)* adalah fungsi kepekatan peluang peubah acak *T*, dan batas atas *–log(1-F(x))* adalah salah satu bentuk batas atas yang dapat digunakan untuk *T* yang terdefinisi di (0, ∞).

Jika *X* suatu peubah acak kontinu maka fungsi kepekatan peluang sebaran *T-X* adalah

$$g(x) = \frac{f(x)}{1-F(x)} r\left(-\log\left(1-F(x)\right)\right) \tag{2}$$

Untuk *T* bersebaran Gamma*(α, β)* maka

$$r(t) = \left(\beta^{\alpha}\Gamma(\alpha)\right)^{-1} t^{\alpha-1} e^{-t/\beta}, t \geq 0, \tag{3}$$

Substitusi (3) ke (2) menghasilkan fungsi kepekatan peluang sebaran gamma-*X*

$$\begin{aligned} g(x) &= \frac{f(x)}{1-F(x)} \left(\beta^{\alpha}\Gamma(\alpha)\right)^{-1} \left(-\log\left(1-F(x)\right)\right)^{\alpha-1} e^{-\left(-\log\left(1-F(x)\right)\right)/\beta} \\ &= \frac{f(x)}{1-F(x)} \left(\beta^{\alpha}\Gamma(\alpha)\right)^{-1} \left(-\log\left(1-F(x)\right)\right)^{\alpha-1} \left(1-F(x)\right)^{1/\beta} \\ &= f(x)\left(\beta^{\alpha}\Gamma(\alpha)\right)^{-1} \left(-\log\left(1-F(x)\right)\right)^{\alpha-1} \left(1-F(x)\right)^{1/\beta-1} \end{aligned} \tag{4}$$

Jika *X* bersebaran Pareto (*θ, k*) dengan fkp

$$f(x) = k\theta^{k} x^{-(k+1)}, x > \theta$$

dan Fungsi sebaran

$$F(x) = 1 - \left(\frac{\theta}{x}\right)^{k}, x > \theta$$

substitusi kedua fungsi terakhir ini ke dalam (4)

$$\begin{aligned} g(x) = k\theta^{k} x^{-(k+1)} \left(\beta^{\alpha}\Gamma(\alpha)\right)^{-1} \left(-\log\left(1-\left(1-\left(\frac{\theta}{x}\right)^{k}\right)\right)\right)^{\alpha-1} \Bigg(1 \\ - \left(-\left(\frac{\theta}{x}\right)^{k}\right)\Bigg)^{1/\beta-1} \end{aligned}$$

$$\begin{aligned} &= \frac{k}{x}\left(\frac{\theta}{x}\right)^{k} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \left(-\log\left(\left(\frac{\theta}{x}\right)^{k}\right)\right)^{\alpha-1} \left(\left(\frac{\theta}{x}\right)^{k}\right)^{1/\beta-1} \\ &= \frac{k}{x}\left(\frac{\theta}{x}\right)^{k} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \left(k\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1} \left(\frac{\theta}{x}\right)^{k/\beta} \left(\frac{\theta}{x}\right)^{-k} \\ &= \frac{k^{\alpha}}{x} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1} \left(\frac{\theta}{x}\right)^{k/\beta} \end{aligned}$$

Dengan mengambil *c=β/k*

$$g(x) = \frac{1}{x}\frac{1}{c^{\alpha}\Gamma(\alpha)} \left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1} \left(\frac{\theta}{x}\right)^{1/c} \tag{5}$$

dengan *α, c, θ > 0 dan x>θ*. Persamaan (5) adalah fkp sebaran Gamma-Pareto.

5 Peluang Penelitian

Perkembangan bentuk sebaran membuka banyak peluang penelitian. Beberapa bentuk penelitian yang dapat dilakukan diuraikan berikut.

ISBN: 978-979-792-552-9

Pembentukan sebaran baru

Aturan yang jelas dalam membentuk sebaran baru dapat dimanfaatkan untuk membentuk sebaran-sebaran baru. Ada beberapa cara membuat sebaran *T-X* yang baru. Cara pertama adalah membentuk keluarga sebaran baru. Cara kedua adalah sebaliknya,yaitu mencari peubah acak *X* yang belum digunakan untuk melengkapi keluarga *T-X* yang sudah ada. Cara ketiga berhubungan dengan fungsi *W(F(x))* yang digunakan sebagai batas atas integral bagi fungsi sebaran *T-X*.

Pembentukan keluarga sebaran *T-X* yang baru masih dimungkinkan karena belum semua peubah acak kontinu digunakan sebagi pembangkit sebaran T-X. Artinya dicari peubah acak *T* yang belum digunakan sebagai pembangkit sebaran *T-X*. Alzaatreh *et al* (2013a) memberikan daftar peubah acak kontinu beserta bentuk fkp sebaran *T-X* nya. Beberapa peubah acak kontinu tidak terdapat dalam daftar tersebut. Salah satunya adalah sebaran normal.

Peluang penelitian untuk cara kedua juga masih sangat terbuka karena *X* dapat berupa peubah acak diskret maupun kontinu. Sampai saat ini belum ada peubah acak diskret yang digunakan sebagai *X* dalam sebaran *T-X*. Berdasarkan daftar dari Alzaatreh *et al* (2013a) tersebut dapat dibuat banyak sebaran *T-X*. Dengan fkp *T-X* yang sudah tersedia, pekerjaan yang ada hanya mengganti *f(x)* dan *F(x)* dengan fkp dan fungsi sebaran *X* yang dipilih. Hasil yang didapat adalah fkp sebaran *T-X* dalam fungsi *X*. Selanjutnya, fkp tersebut diintegralkan untuk mendapatkan fungsi sebarannya.

Tersedianya beberapa pilihan fungsi *W(F(x))* untuk daerah definisi tertentu dari peubah acak T, memungkinkan peneliti untuk memilih satu diantaranya. Misalnya, yang paling sederhana, untuk $T \epsilon (0, 1)$, *W(F(x))* dapat dipilih dari *F(x)* atau $F^{\alpha}(x)$. Dalam hal ini pilihan ditentukan oleh nilai α. Masih perlu diteliti apakah penggunaan nilai α yang berbeda akan menghasilkan sebaran *T-X* yang sama. Jika hasilnya berbeda tentu saja akan banyak sekali sebaran *T-X* yang dapat dibentuk dari sepasang peubah acak *T* dan *X*

Kajian sifat-sifat sebaran yang sudah ada

Kajian mengenai sifat-sifat sebaran yang sudah terbentuk banyak dilakukan peneliti. Sebagai contoh, kajian sifat *bimodality* dari sebaran beta-normal oleh Famoye *et al.*(2004). Banyak sifat sebaran yang perlu diteliti. Sifat-sifat tersebut mungkin berhubungan dengan sebaran data seperti *mean*, median, modus, ragam, kemenjuluran (*skewness*), dan kurtosis. Untuk mengetahui sifat-sifat, kecuali median dan modus, tersebut tentu saja diperlukan bentuk momen ke-k dari sebaran *T-X*. Sifat-sifat lain adalah yang berhubungan dengan pemodelan yaitu sifat-sifat penduga parameter. Hal lain yang perlu juga diteliti adalah kelebihan dan kekurangan sebaran *T-X*. Hal ini perlu terutama untuk membandingkannya dengan peubah asalnya atau bentuk pengembangan lain dari peubah asal tersebut. Misalnya Gamma-Pareto dibandingkan dengan beta-pareto atau pareto terampat.

Pemodelan berdasarkan sebaran yang sudah ada

Diantara sekian banyak sebaran *T-X* yang sudah dibuat, baru beberapa yang dikembangkan pemodelannya. Ortega *et al*(2011) mengembangkan model regresi untuk

ISBN: 978-979-792-552-9

Log-Beta Weibull. Sementara Pascoa *et al* (2013) Log-Kumaraswamy Generalized Gamma. Keduanya menggunakan model skala-lokasi (*location-scale model*).

Bentuk model yang akan dikembangkan tergantung pada bentuk data dan tujuan pemodelan. Model yang paling umum dikembangkan dari suatu sebaran adalah model regresi. Model yang mengkaji pengaruh peubah bebas terhadap peubah respon. Model regresi banyak bentuknya. Untuk peubah acak yang tidak menyebar normal, biasanya dibentuk model linear terampat. Bila peubah bebas ada yang tetap dan ada yang acak, dikembangkan model campuran (*mixed model*). Khusus untuk selain sebaran normal, model campuran berbentuk *nonlinear mixed model*. Tetapi kadang juga diarahkan ke model linear campuran terampat (Generalised linear mixed model, GLMM). Model lain yang juga sering digunakan adalah untuk masalah *lifetime* dengan data tersensor. Jadi ada beberapa pilihan model yang dapat dikembangkan dari sebaran *T-X*.

Banyak hal yang harus dilakukan dalam pengembangan model dari peubah acak. Hal terpenting adalah membentuk modelnya. Selanjutnya diperlukan metode untuk pendugaan parameter model. Metode kemungkinan maksimum merupakan metode yang populer digunakan. Setelah didapat penduga parameter, diperlukan alat untuk memeriksa kebaikan model. Sifat-sifat penduga parameter juga perlu dikaji untuk menjamin bahwa penduga tersebut memenuhi kriteria penduga yang baik bagi parameter model.

Penerapan: kesesuaian sebaran dan model

Penerapan sebaran peubah acak untuk data empirik meliputi 2 bentuk. Pertama adalah pengujian kesesuaian data dengan sebaran tertentu. Bentuk yang kedua penerapan model. Perbedaan keduanya adalah bentuk yang kedua menggunakan peubah bebas untuk menjelaskan peubah respon, sedangkan yang pertama hanya menganalisis bentuk sebaran peubah respon.

Sebagian besar penulis sebaran *T-X* menyertakan terapan untuk sebaran yang mereka buat. Tetapi terapannya hanya terbatas pada kesesuaian data dengan sebaran yang mereka buat atau sebaliknya menunjukkan bahwa sebaran mereka sesuai untuk data jenis tertentu. Istilah *fitting data* biasa digunakan untuk terapan seperti ini. Alzaatreh *et al* (2012) menunjukkan kesesuaian sebaran Gamma-Pareto untuk 3 jenis data. Ketiganya adalah data dengan bentuk J terbalik, data yang hampir simetri, dan data dengan ekor kanan yang panjang. Dalam penerapan untuk kesesuaian data tersebut ada pendugaan parameter sebaran. Ukuran kebaikan suai (*goodness of fit*) seperti *Akaike Information Criteria* (AIC) dijadikan ukuran kebaikan untuk kesesuaian data-sebaran tersebut. Peluang penelitian jenis ini sangat banyak karena banyaknya jenis data dengan bentuk sebaran yang bermacam-macam. Tugas peneliti adalah membuat plot data kemudian mencari sebaran yang mungkin sesuai untuk data tersebut.

Sedikitnya model yang sudah tersedia untuk diterapkan bukan berarti peluang penelitiannya juga sedikit. Sebagai contoh, Ortega *et al*(2011) menerapkan model regresi log-beta weibull untuk menduga kesembuhan dari kanker prostat. Banyak data di dunia nyata yang sejenis dengan data *lifetime* ini. Belum lagi ditambah dengan pilihan peubah bebas yang banyak. Artinya peluang terapan untuk model yang sudah tersedia juga masih besar.

ISBN: 978-979-792-552-9

6 Penutup

Dalam pengembangan sebaran peluang *T-X* masih sangat terbuka peluang penelitian terkini. Peluang tersebut hendaknya dimanfaatkan oleh pengajar, peneliti, dan praktisi di Indonesia agar dapat ikut serta dalam perkembangan teori dan terapan sebaran peluang terutama sebaran *T-X*.

Daftar Pustaka

[1] A Akinsete, F Famoye dan C Lee, The Beta-Pareto Distribution, *Statistics (*2008), 42:6, 547-563. DOI 1080/023318801983876

[2] A Alzaatreh, F Famoye dan F Lee, The Gamma-Normal Distribution: Properties and Application. *Computational statistics and data analysis* 69 (2014), 67-80

[3] A Alzaatreh, F Famoye, dan C Lee, Gamma-Pareto Distribution And Application, *Journal of Modern Applied Statistical Methods*. Vol. 11 (2012) Issue 1, Article 7

[4] A Alzaatreh, F Famoye, dan C Lee, A New Method For Generating Families Of Continous Distributions. *Metron* 71 (2013a): 63-79 DOI 10.1007/s40300-013-0007-y

[5] A Alzaatreh, F Famoye, dan C Lee, Weibull-Pareto Distribution and Its Applications, *Commun in Stat-Theory and Methods* 42 (2013b): 1673-1691 DOI.10.1080/03610926.2011.599002

[6] F Famoye, C Lee, dan N Eugene, "Beta-Normal Distribution: Bimodality Properties and Application," *Journal of Modern Applied Statistical Methods*: Vol. 3(2004): Iss. 1, Article 10. http://digitalcommons.wayne.edu/jmasm/vol3/iss1/10

[7] F Famoye, C Lee, dan O Olumolade, The beta-Weibull distribution, *J. Stat.Theory Appl*.4(2) (2005) : 121-136

[8] GM Cordeiro ,EMM Ortega, dan S Nadarajah, The Kumaraswamy Weibull distribution with application to failure data, *J.Frankl.Inst*. 347 (2010): 1399-1429

[9] L Kong, C Lee, dan JH Sepanski, On theproperties of beta-gamma distribution. *J.Mod.Appl.Stat.Methods* 6(1)(2007):187-211

[10] MC Jones, Kumaraswamys distribution : a beta-type distribution with tractability advantages, *Stat.Methodol*. 6 (2009) : 70-81

[11] N Eugene, C Lee. dan F Famoye, The Beta-Normal Distribution And It Application. *Common.stat.Theory Methods* 31(4) (2002): 497-512

[12] S Nadarajah dan S Kotz, The beta exponential distribution. *Reliab.Eng.Systt.Saf*. 91(6) (2005): 689-697

[13] EMM Ortega, GM Cordeiro dan MW Kattan. 2011. Log-Beta Weibull Regression Model With Application To Predict Recurrence Of Prostate Cancer. *Statistical Paper* Vol 54, Issue1 pp 113-132 DOI 10.1007/s 00362-011-0414-1 Springer-Verlag

[14] MAR Pascoa, CMM de Paiva, GM Cordeiro dan EMM Ortega. 2013. The Log-Kumaraswamy Generalized Gamma Regression Model with Application to Chemical Dependency Data. *Journal of Data Science* 11: 781-818