Quadratic Functional Variable Transformation of the Exponential-Gamma Distribution for Flexible Lifetime Modelling

Taiwo Michael Ayeni[1], **Elijah A. Odukoya**[2], **Yichuan Zhao**[1]

[1]Department of Mathematics & Statistics, Georgia State University, Atlanta, GA

[2]Federal University of Technology & Environmental Sciences, Iyin-Ekiti, Nigeria

NISS Virtual New Researchers Conference 2026

Georgia State University

Real lifetime datasets are heavily right-skewed

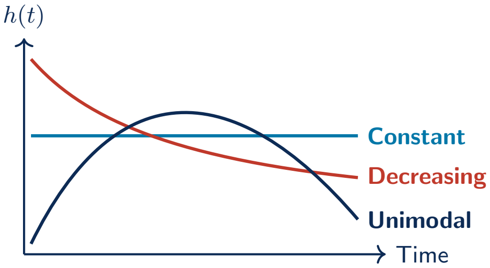
Lifetime data in reliability and biomedical studies often show **early failure**, **random failure**, or **wear-out** patterns, motivating the search for a flexible yet parsimonious lifetime model.

Limitations of Classical Models

Model	Limitation
Exponential	Constant hazard only
Gamma / Weibull	Monotone hazard only
Multi-parameter	Loses tractability

Research Goal

Derive a **tractable two-parameter** lifetime distribution with **closed-form** density, hazard, moments, and entropy, supporting all three hazard regimes from a single parsimonious family.



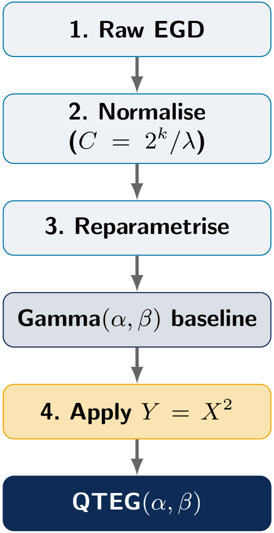
From EGD to QTEG: The Construction Pipeline

1. **Raw EGD** [Ogunwale et al., 2019]
   Multiplies Exponential and Gamma densities, but **does not integrate to 1** (integrates to $\lambda/2^k \neq 1$)

2. **Normalisation**
   Apply $C = 2^k/\lambda$. Omitting C shifts all information criteria by $|2n \log(\lambda/2^k)|$, **invalidating model comparisons**

3. **Reparametrisation** $(\alpha = k,\ \beta = 2\lambda)$
   Corrected baseline $\equiv$ **Gamma**(α, β)

4. **Quadratic transformation** $Y = X^2$
   Gives QTEG with heavier tail decay $e^{-\beta\sqrt{y}}$



Step 2: The Quadratic Transformation $Y = X^2$

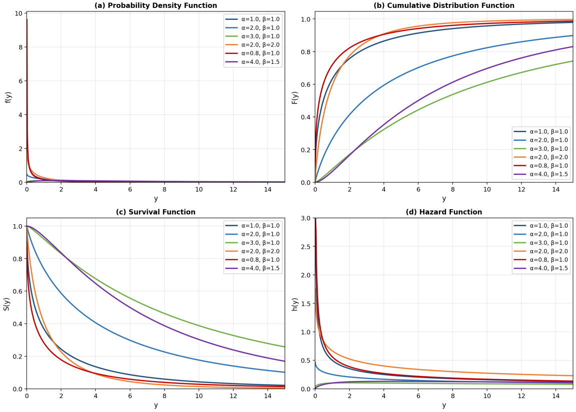
Functional Variable Transformation [Alzaatreh et al., 2013, Casella and Berger, 2002]

For $Y = g(X)$ strictly increasing,
$f_Y(y) = f_X(g^{-1}(y))\,|d\,g^{-1}/dy|$.

Transform	$g(x)$	$g^{-1}(y)$	Tail
Quadratic	x^2	$\sqrt{y}$	$e^{-\beta\sqrt{y}}$
Identity	x	y	$e^{-\beta y}$
Power (q)	x^q	$y^{1/q}$	$e^{-\beta y^{1/q}}$

Three gains from $Y = X^2$

- **Heavier tail:** $e^{-\beta x} \to e^{-\beta\sqrt{y}}$
- **Variance amplified:** no new parameters
- **Hazard regime** via α alone



QTEG PDF, CDF, survival function, and hazard rate

The Density (one key formula)

$$f_Y(y;\alpha,\beta) = \frac{\beta^\alpha}{2\,\Gamma(\alpha)}\, y^{\frac{\alpha-2}{2}}\, e^{-\beta\sqrt{y}}, \quad y > 0$$

What the parameters do

- α (shape): controls hazard regime and skewness, the key modelling lever
- β (rate): pure scale parameter; shifts the distribution left/right without changing shape
- CV is **independent of** β: relative spread depends only on α

All results in closed form

Quantity	Key feature
CDF	Regularised incomplete gamma $I(\alpha, \beta\sqrt{y})$
Hazard rate	Unimodal when $\alpha > 2$; shape via α only
Mean	$\alpha(\alpha+1)/\beta^2$
nth raw moment	$\Gamma(2n+\alpha)/(\beta^{2n}\Gamma(\alpha))$
Shannon entropy	Closed form via digamma $\psi(\alpha)$
Quantile $Q(p)$	$(I^{-1}(\alpha,p)/\beta)^2$
Order statistics	Exact PDF via incomplete gamma

Every quantity needed for inference, reliability analysis, and model comparison is available in **closed form**.

How the MLE works: 4 steps

1. **Score equation for β solves exactly**
   $\hat{\beta}(\alpha) = \alpha \,/\, \sqrt{y}$ (no numerical search needed)
2. **Substitute into α-equation**
   Reduces a 2D optimisation to a **one-dimensional profile search** in α
3. **Profile likelihood is strictly concave**
   Proved analytically via the trigamma inequality, guarantees a **unique global maximum**
4. **Compute standard errors**
   From the analytic Fisher information matrix; invert for 95% Wald confidence intervals

What this gives you

- **Existence & uniqueness** of MLE proved (no convergence ambiguity)
- **Semi-closed starting value** for β at each grid point guarantees global convergence
- **Analytic Fisher information** matrix: standard errors require no bootstrap
- Grid of 8 starting values spans decreasing, constant, and unimodal hazard regimes

100%	93.8–96.4%	$\downarrow O(1/n)$
Convergence all scenarios	Empirical CP nominal 95%	Bias & RMSE as n grows

Design

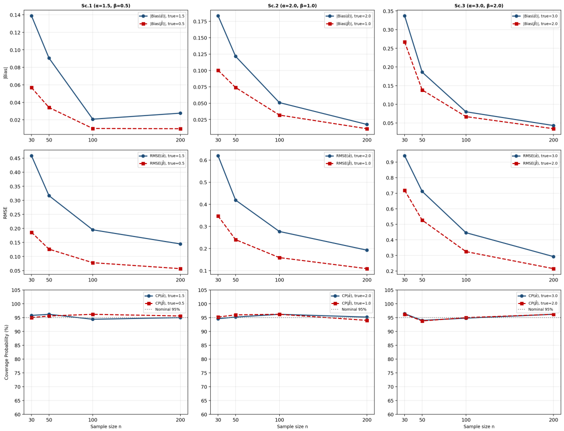
$N = 500$ replications; $n \in \{30, 50, 100, 200\}$

	Hazard	α	β
Sc. 1	Decreasing	1.5	0.5
Sc. 2	Constant	2.0	1.0
Sc. 3	Unimodal	3.0	2.0

Key findings

- **100%** convergence across all 12 scenario–n cells
- Bias & RMSE **decrease monotonically** with n
- Coverage **93.8%–96.4%** (nominal 95%)
- AvgSE $\approx$ RMSE: Hessian is accurate

Figure 2: Monte Carlo Simulation Study (N = 500 replications)



Empirical Applications: Three Benchmark Datasets

Datasets

- **DS1:** Bladder cancer remission, n=128
- **DS2:** Boeing 720 failures, n=213
- **DS3:** Melanoma survival, n=205

All: CV > 1, skewness > 2, kurtosis > 4.

Competing models (8)

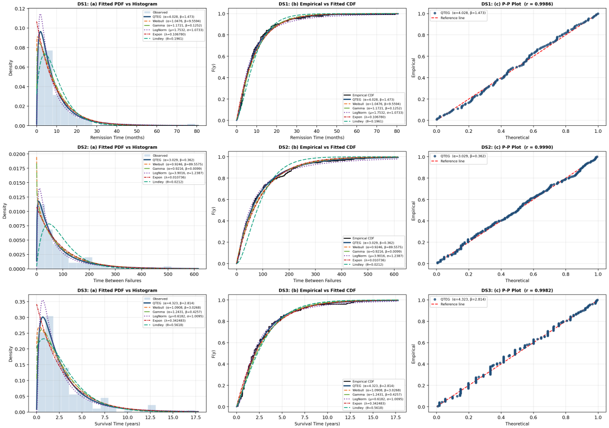
Weibull, Gamma, Log-Normal, Exponential, EGD, Exp-Gamma, Kum-Gamma, Lindley.

Criteria (7)

AIC, BIC, CAIC, HQIC; KS p-val, CvM, AD.

QTEG ranks #1 on all 7 criteria across all 3 datasets.

Figure 3: Empirical Applications -- Fitted PDF, CDF, and P-P Plots
DS1 Bladder Cancer, DS2 Boeing 720 Failures, DS3 Malignant Melanoma



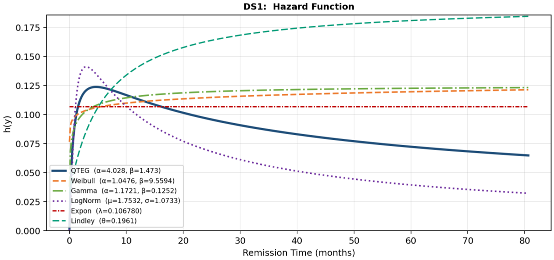
DS1 Bladder Cancer: Full Comparison ($n = 128$)

Model	LogLik	AIC	BIC	CAIC	HQIC	KS p
QTEG	**−410.9**	**825.7**	**831.4**	**825.8**	**828.0**	**0.953**
Exp-Gamma	−413.4	830.7	836.4	830.8	833.1	0.484
Weibull	−414.1	832.2	837.9	832.3	834.5	0.543
Exp-Gamma*	−411.3	828.7	837.2	828.9	832.2	0.807
Kum-Gamma	−411.1	830.3	841.7	830.6	834.9	0.869
LogNormal	−415.1	834.2	839.9	834.3	836.5	0.698
Exponential	−414.3	830.7	833.5	830.7	831.8	0.329

Bold = best; QTEG is **the only model** with KS $p > 0.05$.

All three datasets: QTEG ranks #1 on all 4 information criteria.

Fitted hazard: DS1 Bladder Cancer



QTEG captures the unimodal hazard ($\hat{\alpha} = 4.03$); competitors impose monotone or constant shapes.

Summary of Contributions

Theory

- Normalised the EGD baseline; proved QTEG $\neq$ Gamma
- Closed-form PDF, CDF, hazard, survival, entropy, order statistics, quantile
- Analytic Fisher information matrix
- Proved MLE existence & uniqueness (trigamma inequality)

Simulation & Applications

- Monte Carlo: $N{=}500$, three hazard regimes, $n \in \{30, 50, 100, 200\}$
- 100% convergence; correct 95% coverage
- Three benchmark lifetime datasets
- Ranks #1 on all 7 goodness-of-fit criteria vs. 8 competing models

Status: Under review, *Statistical Papers*, 2026

Bayesian (β has conjugate Gamma posterior) | Censored-data & survival regression | Jackknife EL goodness-of-fit test | Generalised $Y = X^q$, $q > 0$

Acknowledgements

Advisor & Co-authors

The author gratefully thanks **Professor Yichuan Zhao** (Georgia State University) for invaluable guidance, mentorship, and continuous support throughout this research.

Thanks also to co-author **Dr. Elijah A. Odukoya** (Federal University of Technology & Environmental Sciences, Nigeria) for fruitful collaboration on this work.

Institutional & Conference

The author thanks the **Department of Mathematics and Statistics, Georgia State University** for institutional support, and the **NISS New Researchers Network** for the opportunity to present at the Virtual New Researchers Conference 2026.

Thank you | Questions welcome

A. Alzaatreh, C. Lee, and F. Famoye. A new method for generating families of continuous distributions. *METRON*, 71:63–79, 2013. doi: https://doi.org/10.1007/s40300-013-0007-y.

G. Casella and R.L. Berger. *Statistical Inference*. Duxbury, Pacific Grove, CA, 2nd edition, 2002.

O.D. Ogunwale, O.A. Adewusi, and T.M. Ayeni. Exponential-Gamma distribution. *International Journal of Emerging Technology and Advanced Engineering*, 9(10), 2019.