Quiz #1, Survival Analysis I, 2017 Spring [+22 points] +21
Name: 張雯婧

- **Not only answer but also derivations for every question**

+4 **Q1[+4]** The p.d.f. of the generalized gamma distribution is defined as

$f(x) = Cx^{\alpha\beta-1}e^{-\lambda x^\alpha}$, where C is a constant, $x > 0, \alpha > 0, \beta > 0$, and $\lambda > 0$. $0<x<\infty \Rightarrow 0<t<\infty$

$\frac{\alpha\lambda^\beta}{\Gamma(\beta)}$ $x^{\alpha\beta} \cdot x^{-1}$ $\Gamma(\alpha\beta, \frac{1}{\lambda}$

+1 (i) [+1] Derive C. Let $t=\lambda x^\alpha \Rightarrow dt = \lambda\alpha x^{\alpha-1}dx \Rightarrow x = (\frac{t}{\lambda})^{\frac{1}{\alpha}}$

$$C\int_0^\infty x^{\alpha\beta-1}e^{-\lambda x^\alpha}dx = 1 \Rightarrow C\int_0^\infty x^{\alpha\beta-\alpha+\alpha-1}e^{-\lambda x^\alpha}dx = 1 \Rightarrow C\cdot\frac{1}{\lambda\alpha}\int_0^\infty (\frac{t}{\lambda})^{\beta-1}e^{-t}dt = 1$$

$$\Rightarrow C\cdot\frac{1}{\alpha\lambda^\beta}\underbrace{\int_0^\infty t^{\beta-1}e^{-t}dt}_{=\Gamma(\beta)} = 1 \Rightarrow C = \frac{\alpha\lambda^\beta}{\Gamma(\beta)} \quad \#$$

$S(x) = 1-F(x)$ +2 (ii) [+2] Derive the survival function $S(x)$ using the incomplete Gamma

$=1-\int$

function $I(x;\beta) = \frac{1}{\Gamma(\beta)}\int_0^x t^{\beta-1}e^{-t}dt$. $1-$

$x<y<\infty \Rightarrow \lambda x^\alpha < t < \infty$

Let $t=\lambda y^\alpha \Rightarrow dt = \lambda\alpha y^{\alpha-1}dy \Rightarrow y = (\frac{t}{\lambda})^{\frac{1}{\alpha}}$

$$S(x) = P(X>x) = \frac{\alpha\lambda^\beta}{\Gamma(\beta)}\int_x^\infty y^{\alpha\beta-1}e^{-\lambda y^\alpha}dy = \frac{1}{\Gamma(\beta)}\int_{\lambda x^\alpha}^\infty t^{\beta-1}e^{-t}dt$$

$$= \frac{1}{\Gamma(\beta)}\left[\int_0^\infty t^{\beta-1}e^{-t}dt - \int_0^{\lambda x^\alpha} t^{\beta-1}e^{-t}dt\right] = 1 - I(\lambda x^\alpha, \beta) \quad \#$$

+1 (iii) [+1] Show that the generalized gamma (α, β, λ) distribution reduces to the exponential $(1,1,\lambda)$, $(\alpha,1,\lambda)$ Weibull, and gamma distributions. $(1, \beta, \lambda)$

$X \sim$ G-gamma(α,β,λ) : $f(x) = \frac{\alpha\lambda^\beta}{\Gamma(\beta)}x^{\alpha\beta-1}e^{-\lambda x^\alpha}$, $\alpha>0, \beta>0, \lambda>0, x\geq 0$

If $\alpha=\beta=1$, $X \sim$ Exp (λ) : $f(x) = \lambda e^{-\lambda x}$, $\lambda>0, x\geq 0$

If $\beta=1$, $X \sim$ Weibull (α,λ) : $f(x) = \alpha\lambda x^{\alpha-1}e^{-\lambda x^\alpha}$, $\alpha>0, \lambda>0, x\geq 0$

If $\alpha=1$, $X \sim$ gamma(β,λ) : $f(x) = \frac{\lambda^\beta}{\Gamma(\beta)}x^{\beta-1}e^{-\lambda x}$, $\beta>0, x\geq 0$ #

θx^{-1} $\quad -\theta x^{-2}$

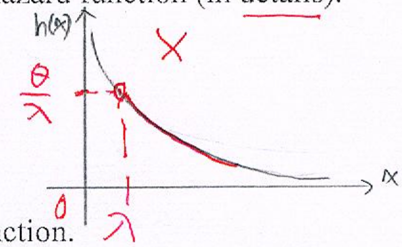
+1

Q2.[+2] The hazard function of the Pareto distribution is defined as $h(x)=\theta/x, x>\lambda, \theta>0$. $\lambda>0$

+0 (i) [+1] Draw the graph of the hazard function (in details).

$h'(x) = \frac{-\theta}{x^2} < 0$ ↘

$h''(x) = \frac{2\theta}{x^3} > 0$



$h(0) \to \infty$

$h(\infty) \to 0$

+1 (ii) [+1] Derive the survival function.

$H(x) = \int_\lambda^x \frac{\theta}{x}\,dx = \theta \log x \Big|_\lambda^x = \theta \cdot \log\frac{x}{\lambda}$

$S(x) = \exp(-H(x)) = e^{-\theta\cdot\log(\frac{x}{\lambda})} = \left(\frac{\lambda}{x}\right)^\theta$ #

+3

Q3. [+3] X has a survival function $S(x)=\exp(-\lambda x^\alpha)$, $x, \alpha, \lambda > 0$.

+1 (i) [+1] Derive rth moment $E[X^r]$

$f(x) = -\frac{d}{dx}S(x) = \alpha\lambda x^{\alpha-1} e^{-\lambda x^\alpha}$ $\quad$ * Let $t=\lambda x^\alpha \Rightarrow dt = \alpha\lambda x^{\alpha-1}dx \Rightarrow x = \left(\frac{t}{\lambda}\right)^{\frac{1}{\alpha}}$

$E(X^r) = \alpha\lambda\int_0^\infty x^r x^{\alpha-1} e^{-\lambda x^\alpha}dx = \int_0^\infty \left(\frac{t}{\lambda}\right)^{\frac{r}{\alpha}} e^{-t}dt = \lambda^{-\frac{r}{\alpha}}\int_0^\infty t^{\frac{r}{\alpha}+1-1}e^{-t}dt = \lambda^{-\frac{r}{\alpha}}\cdot\Gamma\left(\frac{r}{\alpha}+1\right)$ #

+1 (ii) [+1] Derive $Var[X]$

$Var(X) = E(X^2) - [EX]^2 = \lambda^{-\frac{2}{\alpha}}\Gamma\left(\frac{2}{\alpha}+1\right) - \lambda^{-\frac{2}{\alpha}}\cdot\Gamma^2\left(\frac{1}{\alpha}+1\right)$ #

+1 (iii) [+1] Write the p.d.f of $Y=\log X$ as a location-scale family by defining parameters (μ, σ). $X \sim Weibull(\alpha, \lambda)$

$e^{\log\lambda}$ ↑

$S(y) = P(Y>y) = P(\log X > y) = P(X > e^y) = \exp[-\lambda e^{\alpha y}] = \exp[-e^{\alpha y + \log\lambda}] = \exp\left[-e^{\alpha\left[y - \left(\frac{-\log\lambda}{\alpha}\right)\right]}\right]$

$= \exp\left[-e^{\frac{y-u}{\sigma}}\right]$, where $u = \frac{-\log\lambda}{\alpha}$, $\sigma = \frac{1}{\alpha}$

$f(y) = -\frac{d}{dy}S(y) = \frac{1}{\sigma}e^{\frac{y-u}{\sigma}}\exp\left[-e^{\frac{y-u}{\sigma}}\right] = \frac{1}{\sigma}\exp\left[\frac{y-u}{\sigma} - e^{\frac{y-u}{\sigma}}\right]$, Let $\phi(w) = \exp[w - e^w]$

$= \frac{1}{\sigma}\phi\left(\frac{y-u}{\sigma}\right)$, where $u = \frac{-\log\lambda}{\alpha}$, $\sigma = \frac{1}{\alpha}$ is a location-scale family #

$\sim EV(u, \sigma^2)$

$\left[e^{\theta\lambda_1 t_1^{\alpha_1}}+e^{\theta\lambda_2 t_2^{\alpha_2}}-1\right]^{-\frac{1}{\theta}}$

+4 **Q4 [+4]** A pair of random variables (T_1, T_2) follows the joint survival function

$$S_\theta(t_1,t_2)=\Pr(T_1>t_1,T_2>t_2)=\left[\exp(\theta\lambda_1 t_1^{\alpha_1})+\exp(\theta\lambda_2 t_2^{\alpha_2})-1\right]^{-\frac{1}{\theta}},$$

$$t_1,t_2\ge 0,\quad \alpha_1,\alpha_2>0,\quad \theta>0.$$

+1 (i) [+1] Derive the marginal survival function and the marginal hazard function for T_1.

$S_1(t)=S(t,0)=e^{\theta\lambda_1 t^{\alpha_1}\cdot(-\frac{1}{\theta})}=e^{-\lambda_1 t^{\alpha_1}}$ ✓

$S_2(t)=S(0,t)=e^{-\lambda_2 t^{\alpha_2}}$ #

$h_1(t)=\dfrac{-\frac{d}{dt}S_1(t)}{S_1(t)}=\dfrac{-(-\alpha_1\lambda_1 t^{\alpha_1-1})e^{-\lambda_1 t^{\alpha_1}}}{e^{-\lambda_1 t^{\alpha_1}}}=\alpha_1\lambda_1 t^{\alpha_1-1}$ # ✓

+1 (ii) [+1] Derive the cause-specific hazard function $h_1^{\#}(t)$.

$h_1^{\#}(t)=\dfrac{-\frac{d}{dt_1}S(t_1,t_2)\big|_{t_1=t_2=t}}{S(t,t)}=\dfrac{\cancel{\frac{1}{\theta}}\left[e^{\theta\lambda_1 t_1^{\alpha_1}}+e^{\theta\lambda_2 t_2^{\alpha_2}}-1\right]^{-\frac{1}{\theta}-1}\cdot\cancel{\theta}\alpha_1\lambda_1 t_1^{\alpha_1-1}e^{\theta\lambda_1 t_1^{\alpha_1}}\Big|_{t_1=t_2=t}}{\left[e^{\theta\lambda_1 t^{\alpha_1}}+e^{\theta\lambda_2 t^{\alpha_2}}-1\right]^{-\frac{1}{\theta}}}$

$=\dfrac{\alpha_1\lambda_1 t^{\alpha_1-1}e^{\theta\lambda_1 t^{\alpha_1}}}{e^{\theta\lambda_1 t^{\alpha_1}}+e^{\theta\lambda_2 t^{\alpha_2}}-1}$ # ✓

+1 (iii) [+1] Derive the hazard function for $T=\min(T_1,T_2)$ by calculating

$$h_T(t)=-\frac{dS_T(t)/dt}{S_T(t)},\text{ where }S_T(t)=\Pr(T\ge t).$$

$*\ S_T(t)=P(T\ge t)=P(\min(T_1,T_2)\ge t)=P(T_1\ge t,T_2\ge t)=S(t,t)$

$h_T(t)=\dfrac{-\frac{d}{dt}S(t,t)}{S(t,t)}=\dfrac{\cancel{\frac{1}{\theta}}\left[e^{\theta\lambda_1 t^{\alpha_1}}+e^{\theta\lambda_2 t^{\alpha_2}}-1\right]^{-\frac{1}{\theta}-1}\left[\cancel{\theta}\alpha_1\lambda_1 t^{\alpha_1-1}e^{\theta\lambda_1 t^{\alpha_1}}+\cancel{\theta}\alpha_2\lambda_2 t^{\alpha_2-1}e^{\theta\lambda_2 t^{\alpha_2}}\right]}{\left[e^{\theta\lambda_1 t^{\alpha_1}}+e^{\theta\lambda_2 t^{\alpha_2}}-1\right]^{-\frac{1}{\theta}}}$

$=\dfrac{\alpha_1\lambda_1 t^{\alpha_1-1}e^{\theta\lambda_1 t^{\alpha_1}}+\alpha_2\lambda_2 t^{\alpha_2-1}e^{\theta\lambda_2 t^{\alpha_2}}}{e^{\theta\lambda_1 t^{\alpha_1}}+e^{\theta\lambda_2 t^{\alpha_2}}-1}$ # ✓

+1 (iv) [+1] Derive $h_1^{\#}(t)+h_2^{\#}(t)$.

by Q4 (ii) $h_1^{\#}(t)=\dfrac{\alpha_1\lambda_1 t^{\alpha_1-1}e^{\theta\lambda_1 t^{\alpha_1}}}{e^{\theta\lambda_1 t^{\alpha_1}}+e^{\theta\lambda_2 t^{\alpha_2}}-1}$

$h_2^{\#}(t)=\dfrac{-\frac{d}{dt_2}S(t_1,t_2)\big|_{t_1=t_2=t}}{S(t,t)}=\dfrac{\cancel{\frac{1}{\theta}}\left[e^{\theta\lambda_1 t_1^{\alpha_1}}+e^{\theta\lambda_2 t_2^{\alpha_2}}-1\right]^{-\frac{1}{\theta}-1}\cdot\cancel{\theta}\alpha_2\lambda_2 t_2^{\alpha_2-1}\cdot e^{\theta\lambda_2 t_2^{\alpha_2}}\Big|_{t_1=t_2=t}}{\left[e^{\theta\lambda_1 t^{\alpha_1}}+e^{\theta\lambda_2 t^{\alpha_2}}-1\right]^{-\frac{1}{\theta}}}=\dfrac{\alpha_2\lambda_2 t^{\alpha_2-1}e^{\theta\lambda_2 t^{\alpha_2}}}{e^{\theta\lambda_1 t^{\alpha_1}}+e^{\theta\lambda_2 t^{\alpha_2}}-1}$

$\therefore h_1^{\#}(t)+h_2^{\#}(t)=\dfrac{\alpha_1\lambda_1 t^{\alpha_1-1}e^{\theta\lambda_1 t^{\alpha_1}}+\alpha_2\lambda_2 t^{\alpha_2-1}e^{\theta\lambda_2 t^{\alpha_2}}}{e^{\theta\lambda_1 t^{\alpha_1}}+e^{\theta\lambda_2 t^{\alpha_2}}-1}=h_T(t)$ # ✓

+6

Q5 [+6]. Time-to-death due to K different causes are denoted by $(X_1,\ldots,X_K)$. Time to death is denoted by $T=\min(X_1,\ldots,X_K)$ and the cause of death is denoted by $\delta\in\{1,\ldots,K\}$.

+1 (i) [+1] Write the definition of δ in terms of $(X_1,\ldots,X_K)$ and T.

$$\delta=\begin{cases}1, & \text{if } T=X_1\\ 2, & \text{if } T=X_2\\ \vdots & \\ K, & \text{if } T=X_K\end{cases} \quad \text{where } P(\delta=j)=P(T=X_j)$$

$$\sum_{j=1}^{k} P(\delta=j)=1$$

+1 (ii) [+1] Write the definition of the cause-specific hazard $h_j^{\#}(t)$ in terms of T and δ.

$$h_j^{\#}(t)=\lim_{\Delta t\to 0}\frac{P(t\le T<t+\Delta t,\ \delta=j\mid T\ge t)}{\Delta t}\quad\left(=\lim_{\Delta t\to 0}\frac{P(t\le X_j<t+\Delta t,\ T=X_j\mid X_j\ge t, j=1,\cdots,k)}{\Delta t}\right)$$ #

+1 (iii) [+1] Under competing risks, what we can observe are (t_i, δ_i)

T δ

+1 (iv) [+1] Why we cannot observe $(X_1,\ldots,X_K)$? Explain by a real example with $K=3$. ex.

$$\delta=\begin{cases}1, & \text{if person is death for lung cancer} \quad T=X_1\\ 2, & \text{if person is death for heart attake.} \quad T=X_2\\ 3, & \text{if person is death for other reason.} \quad T=X_3\end{cases}$$

So, we just can observe one of (X_1, X_2, X_3), which is $T=\min(X_1,X_2,X_3)$

+1 (v) [+1] Let $(X_1,\ldots,X_K)$ be independent random variables with $P(X_j>x)=S_j(x)$, $h_j(x)=-\dfrac{d}{dx}\log S_j(x)$, $j=1,\ldots,K$. Calculate the cause-specific hazard $h_j^{\#}(x)$, $j=1,\ldots,K$.

$$S(t_1,\cdots,t_k)=P(X_1>t_1,\cdots,X_k>t_k)=P(X_1>t_1)\cdots P(X_k>t_k)=\prod_{j=1}^{k}S_j(t_j)$$

$$h_j^{\#}(x)=\frac{-\frac{d}{dt_j}S(t_1,\cdots,t_k)\big|_{t_1=\cdots=t_k=t}}{S(t,\cdots,t)}=\frac{-\frac{d}{dt_j}\prod_{j=1}^{k}S_j(t_j)\big|_{t_1=\cdots=t_k=t}}{\prod_{j=1}^{k}S_j(t)}=\frac{-S_j'(t_j)\cdot\prod_{w\ne j}S_w(t_w)\big|_{t_1=\cdots=t_k=t}}{S_j(t)\cdot\prod_{w\ne j}S_w(t)}$$

$$=\frac{-\frac{d}{dt}S_j(t)}{S_j(t)}=\frac{-1}{S_j(t)}\cdot\frac{d}{dt}S_j(t)=\frac{-d}{dt}\log S_j(t)=h_j(x)$$ #

+1 (vi) [+1] Calculate the hazard for $T=\min(X_1,\ldots,X_K)$

$$h_T(t)=\lim_{\Delta t\to 0}\frac{\sum_{j=1}^{k}P(t\le T<t+\Delta t,\ \delta=j\mid T\ge t)}{\Delta t}=\sum_{j=1}^{k}\lim_{\Delta t\to 0}\frac{P(t\le T<t+\Delta t,\ \delta=j\mid T\ge t)}{\Delta t}=\sum_{j=1}^{k}h_j^{\#}(t)$$

$$=\sum_{j=1}^{k}h_j(t)$$ #

$$\left[e^{\theta\lambda_1 x^{\alpha_1}} + e^{\theta\lambda_2 c^{\alpha_2}} - 1\right]^{-\frac{1}{\theta}}$$

+3

Q6 [+3]. Let X be time-to-death and C_r be censoring time. Assume the model

$$S(x,c) = \Pr(X > x, C_r > c) = \left[\exp(\theta\lambda_1 x^{\alpha_1}) + \exp(\theta\lambda_2 c^{\alpha_2}) - 1\right]^{-\frac{1}{\theta}},$$

$$x, c \geq 0,\ \ \alpha_1, \alpha_2, \theta > 0,\ \ \lambda_1, \lambda_2 > 0.$$

Let (t_i, δ_i), $i = 1,2,\ldots,n$, be censored data, where $t_i = \min(X_i, C_{ri})$ and $\delta_i = I(X_i \leq C_{ri})$. Then, the log-likelihood is written as

$$\log L = n\times(\log\alpha_1\lambda_1 - \log\alpha_2\lambda_2) + r\times(\log\alpha_2\lambda_2)$$

$$+\sum_{i=1}^{n}\delta_i\left[(\alpha_1-1)\cdot\log t_i + \theta\lambda_1 t_i^{\alpha_1} - \log\left(e^{\theta\lambda_1 t_i^{\alpha_1}} + e^{\theta\lambda_2 t_i^{\alpha_2}} - 1\right)\right]$$

$$+\sum_{i=1}^{n}(1-\delta_i)\left[(\alpha_2-1)\cdot\log t_i + \theta\lambda_2 t_i^{\alpha_2} - \log\left(e^{\theta\lambda_1 t_i^{\alpha_1}} + e^{\theta\lambda_2 t_i^{\alpha_2}} - 1\right)\right]$$

$$-\frac{1}{\theta}\sum_{i=1}^{n}\left[\log\left[e^{\theta\lambda_1 t_i^{\alpha_1}} + e^{\theta\lambda_2 t_i^{\alpha_2}} - 1\right]\right]$$

where $r = \sum_{i=1}^{n}\delta_i$.

$$* \ L = \prod_{i=1}^{n} h_1^{\#}(t_i)^{\delta_i}\, h_2^{\#}(t_i)^{1-\delta_i}\cdot S(t_i, t_i), \text{ where } h_j^{\#}(t) = \frac{-\frac{d}{dt_j}S(t_1,t_2)\Big|_{t_1=t_2=t}}{S(t,t)}$$

$$h_1^{\#}(t) = \frac{-\frac{d}{dx}S(x,c)\Big|_{x=c=t}}{S(t,t)} = \frac{\frac{1}{\theta}\left[e^{\theta\lambda_1 x^{\alpha_1}} + e^{\theta\lambda_2 c^{\alpha_2}} - 1\right]^{-\frac{1}{\theta}-1}\cdot\theta\alpha_1\lambda_1 x^{\alpha_1-1}\cdot e^{\theta\lambda_1 x^{\alpha_1}}\Big|_{x=c=t}}{\left[e^{\theta\lambda_1 t^{\alpha_1}} + e^{\theta\lambda_2 t^{\alpha_2}} - 1\right]^{-\frac{1}{\theta}}} = \frac{\alpha_1\lambda_1 t^{\alpha_1-1} e^{\theta\lambda_1 t^{\alpha_1}}}{e^{\theta\lambda_1 t^{\alpha_1}} + e^{\theta\lambda_2 t^{\alpha_2}} - 1}$$

$$h_2^{\#}(t) = \frac{\alpha_2\lambda_2 t^{\alpha_2-1} e^{\theta\lambda_2 t^{\alpha_2}}}{e^{\theta\lambda_1 t^{\alpha_1}} + e^{\theta\lambda_2 t^{\alpha_2}} - 1}$$

$$\log h_1^{\#}(t) = \log\alpha_1\lambda_1 + (\alpha_1-1)\log t + \theta\lambda_1 t^{\alpha_1} - \log\left[e^{\theta\lambda_1 t^{\alpha_1}} + e^{\theta\lambda_2 t^{\alpha_2}} - 1\right]$$

$$\log h_2^{\#}(t) = \log\alpha_2\lambda_2 + (\alpha_2-1)\log t + \theta\lambda_2 t^{\alpha_2} - \log\left[e^{\theta\lambda_1 t^{\alpha_1}} + e^{\theta\lambda_2 t^{\alpha_2}} - 1\right]$$

$$\log S(t,t) = -\frac{1}{\theta}\log\left[e^{\theta\lambda_1 t^{\alpha_1}} + e^{\theta\lambda_2 t^{\alpha_2}} - 1\right]$$

$$\therefore \log L = \boxed{\sum_{i=1}^{n}\delta_i\cdot\log\alpha_1\lambda_1} + \overbrace{\sum_{i=1}^{n}(1-\delta_i)}^{n-\Sigma\delta_i}\cdot\log\alpha_2\lambda_2 + \sum_{i=1}^{n}\delta_i\cdot\left[(\alpha_1-1)\cdot\log t_i + \theta\lambda_1 t_i^{\alpha_1} - \log\left(e^{\theta\lambda_1 t_i^{\alpha_1}} + e^{\theta\lambda_2 t_i^{\alpha_2}} - 1\right)\right]$$

$$+\sum_{i=1}^{n}(1-\delta_i)\left[(\alpha_2-1)\cdot\log t_i + \theta\lambda_2 t_i^{\alpha_2} - \log\left(e^{\theta\lambda_1 t_i^{\alpha_1}} + e^{\theta\lambda_2 t_i^{\alpha_2}} - 1\right)\right] - \frac{1}{\theta}\sum_{i=1}^{n}\log\left[e^{\theta\lambda_1 t_i^{\alpha_1}} + e^{\theta\lambda_2 t_i^{\alpha_2}} - 1\right]$$

$$r\times(\log\alpha_1\lambda_1 - \log\alpha_2\lambda_2)$$

$$= \boxed{\sum_{i=1}^{n}\delta_i(\log\alpha_1\lambda_1 - \log\alpha_2\lambda_2)} + n\cdot(\log\alpha_2\lambda_2) + \sum_{i=1}^{n}\delta_i\left[(\alpha_1-1)\cdot\log t_i + \theta\lambda_1 t_i^{\alpha_1} - \log\left(e^{\theta\lambda_1 t_i^{\alpha_1}} + e^{\theta\lambda_2 t_i^{\alpha_2}} - 1\right)\right]$$

OK

$$+\sum_{i=1}^{n}(1-\delta_i)\left[(\alpha_2-1)\cdot\log t_i + \theta\lambda_2 t_i^{\alpha_2} - \log\left(e^{\theta\lambda_1 t_i^{\alpha_1}} + e^{\theta\lambda_2 t_i^{\alpha_2}} - 1\right)\right] - \frac{1}{\theta}\sum_{i=1}^{n}\log\left[e^{\theta\lambda_1 t_i^{\alpha_1}} + e^{\theta\lambda_2 t_i^{\alpha_2}} - 1\right] \quad \#$$