Question #1:

**Solution 1.** *For part (a):*

*We seek to prove equation* (2.4.3) *from textbook, which is the following eqution:*

*Given continuous positive random variable $X$, the variance is*

$$Var(X) = 2\int_0^\infty tS(t)dt - \Big[\int_0^\infty S(t)dt\Big]^2$$

*Here is the proof:*

*Since we know that*

$$Var(X) = E[X^2] - \Big(E[X]\Big)^2$$

*We compute $E[X^2]$ and $E[X]$ separately.*

*We havve:*

$$\begin{aligned} E[X] &= \int_0^\infty tf(t)dt = -\int_0^\infty tdS(t) = -tS(t)\Big|_0^\infty + \int_0^\infty S(t)dt \\ &= \int_0^\infty S(t)dt \end{aligned}$$

*and*

$$\begin{aligned} E[X^2] &= \int_0^\infty t^2 f(t)dt = -\int_0^\infty t^2 dS(t) = -t^2S(t)\Big|_0^\infty + \int_0^\infty 2tS(t)dt \\ &= 2\int_0^\infty tS(t)dt \end{aligned}$$

*Plug these into the equation for variance, we got*

$$\begin{aligned} Var(X) &= E[X^2] - \Big(E[X]\Big)^2 \\ &= 2\int_0^\infty tS(t)dt - \Big[\int_0^\infty S(t)dt\Big]^2 \end{aligned}$$

*For part (b):*

*We are going to prove the 'lack of memory ' property for exponential distribution:*

$$P(X \geq x+z|X \geq x) = P(X \geq z)$$

*We have:*

$$\begin{aligned} P(X \geq x+z|X \geq x) &= \frac{P(X \geq x+z, X \geq x)}{P(X \geq x)} = \frac{P(X \geq x+z)}{P(X \geq x)} = \frac{\int_{x+z}^\infty \lambda e^{-\lambda t}dt}{\int_x^\infty \lambda e^{-\lambda t}dt} \\ &= \frac{-e^{-\lambda t}\Big|_{x+z}^\infty}{-e^{-\lambda t}\Big|_x^\infty} = \frac{\lambda e^{-\lambda(x+z)}}{\lambda e^{-\lambda x}} \\ &= e^{-\lambda z} = P(X \geq z) \end{aligned}$$

*Also we can compute the mean residual life time for exponential distribution:*

$$\begin{aligned} mrl(x) = E[X - x|X > x] &= \frac{\int_x^\infty (t-x)f(t)dt}{S(x)} = \frac{\int_x^\infty S(t)dt}{S(x)} \\ &= \frac{\int_x^\infty e^{-\lambda t}dt}{e^{-\lambda x}} = \frac{\frac{-1}{\lambda}\int_x^\infty e^{-\lambda t}d(-\lambda t)}{e^{-\lambda x}} \\ &= \frac{-\frac{1}{\lambda}e^{-\lambda t}\Big|_x^\infty}{e^{-\lambda x}} = \frac{\frac{1}{\lambda}e^{-\lambda x}}{e^{-\lambda x}} \\ &= \frac{1}{\lambda} \end{aligned}$$

*For part (c):*

*We show an alternative expression for $mrl(x)$:*

$$\begin{aligned} mrl(x) &= \frac{\int_x^\infty (u-x)f(u)du}{S(x)} = \frac{\int_x^\infty uf(u)du - x\int_x^\infty f(u)du}{S(x)} = \frac{\int_x^\infty uf(u)du - xS(x)}{S(x)} \\ &= \frac{\int_x^\infty uf(u)du}{S(x)} - x \end{aligned}$$

*For part (d):*

*The multiplicative hazard model is:*

$$h(x|\mathbf{Z}) = h_0(x)G(\boldsymbol{\beta}'\mathbf{Z})$$

*We want to show that*

$$S(x|\mathbf{Z}) = [S_0(x)]^{G(\boldsymbol{\beta}'\mathbf{Z})}$$

*We use two different approaches to prove this:*

*The first approach is the same as the one on the textbook in the practical section, we have:*

$$\begin{aligned} S(x|\mathbf{Z}) &= \exp\Big\{-\int_0^x h(u|\mathbf{Z})du\Big\} = \exp\Big\{-\int_0^x h_0(u)G(\boldsymbol{\beta}'\mathbf{Z})du\Big\} \\ &= \exp\Big\{-G(\boldsymbol{\beta}'\mathbf{Z})\int_0^x h_0(u)du\Big\} = \Big(\exp\Big\{-\int_0^x h_0(u))du\Big\}\Big)^{G(\boldsymbol{\beta}'\mathbf{Z})} \\ &= S_0(x)^{G(\boldsymbol{\beta}'\mathbf{Z})} \end{aligned}$$

*The second approach is essentially the same idea using the ODE(ordinary differential equation) theory, but in a different algebraic flavor:*

*Since we have*

$$h(x|\mathbf{Z}) = h_0(x)G(\boldsymbol{\beta}'\mathbf{Z})$$

*and $h(x|\mathbf{Z} = 0) = h_0(x)$, we got $G(0) = 1$.*

*From the equation above we got:*

$$h(x|\mathbf{Z}) = h_0(x)G(\boldsymbol{\beta}'\mathbf{Z}) \Longrightarrow \frac{f(x|\mathbf{Z})}{S(x|\mathbf{Z})} = \frac{f_0(x)}{S_0(x)}G(\boldsymbol{\beta}'\mathbf{Z})$$
$$\Longrightarrow -\frac{d}{dx}\log S(x|\mathbf{Z}) = -G(\boldsymbol{\beta}'\mathbf{Z})\frac{d}{dx}\log S_0(x)$$

*Integrate on both sides, we got:*

$$-\log S(x|\mathbf{Z}) = -G(\boldsymbol{\beta}'\mathbf{Z})\log S_0(x) + Constant$$
$$\Longrightarrow S(x|\mathbf{Z}) = Constant \cdot S_0(x)^{G(\boldsymbol{\beta}'\mathbf{Z})}$$

*Since when* $\mathbf{Z} = \mathbf{0}$ *we have* $G(\boldsymbol{\beta}'\mathbf{Z}) = G(0) = 1$ *as explained before, and also* $S(x|\mathbf{Z} = \mathbf{0}) = S_0(x)$, *we got*

$$S(x|\mathbf{Z} = \mathbf{0}) = S_0(x) = Constant \cdot S_0(x)$$

*which implies* $Constant = 1$.*So we conclude*

$$S(x|\mathbf{Z}) = S_0(x)^{G(\boldsymbol{\beta}'\mathbf{Z})}$$

Question #2:

**Solution 2.** *For part (a):*

*(i) If* $X \sim Weibull(\lambda, \alpha)$, *we have the pdf of* $X$ *as*

$$f(x) = \alpha\lambda x^{\alpha-1}\exp(-\lambda x^{\alpha})$$

*with survival function*

$$S(x) = \exp[-\lambda x^{\alpha}]$$

*So the mean for weibull distribution is*

$$\begin{aligned}
E[X] &= \int_0^{\infty} x \cdot \alpha\lambda x^{\alpha-1}\exp(-\lambda x^{\alpha})dx = \int_0^{\infty} \alpha\lambda x^{\alpha} e^{-\lambda x^{\alpha}} dx \\
&= \int_0^{\infty} x \cdot e^{-\lambda x^{\alpha}} d\lambda x^{\alpha} \\
&= \frac{1}{\lambda^{1/\alpha}} \int_0^{\infty} (\lambda x^{\alpha})^{(1+\frac{1}{\alpha})-1} e^{-\lambda x^{\alpha}} d\lambda x^{\alpha} \\
&= \frac{\Gamma(1+\frac{1}{\alpha})}{\lambda^{1/\alpha}}
\end{aligned}$$

*To compute the variance, we first compute the second moment:*

$$\begin{aligned}
E[X^2] &= \int_0^{\infty} x^2 \alpha\lambda x^{\alpha-1} e^{-\lambda x^{\alpha}} dx = \int_0^{\infty} x^2 e^{-\lambda x^{\alpha}} d\lambda x^{\alpha} \\
&= \frac{1}{\lambda^{2/\alpha}} \int_0^{\infty} (\lambda x^{\alpha})^{\frac{2}{\alpha}} e^{-\lambda x^{\alpha}} d\lambda x^{\alpha} = \frac{1}{\lambda^{2/\alpha}} \int_0^{+\infty} (\lambda x^{\alpha})^{(\frac{2}{\alpha}+1)-1} e^{-\lambda x^{\alpha}} d\lambda x^{\alpha} \\
&= \frac{\Gamma(\frac{2}{\alpha}+1)}{\lambda^{\frac{2}{\alpha}}}
\end{aligned}$$

*So the variance of weibull distribution is:*

$$
\begin{aligned}
Var[X] = E[X^2] - \Big(E[X]\Big)^2 &= \frac{\Gamma(\frac{2}{\alpha}+1)}{\lambda^{\frac{2}{\alpha}}} - \Big(\frac{\Gamma(1+\frac{1}{\alpha})}{\lambda^{1/\alpha}}\Big)^2 \\
&= \frac{\Big(\Gamma(\frac{2}{\alpha}+1) - \Gamma^2(1+\frac{1}{\alpha})\Big)}{\lambda^{\frac{2}{\alpha}}}
\end{aligned}
$$

*(ii) If $X \sim \Gamma(\beta, \lambda)$(I am following the notation from textbook, see Table 2.2 on page 38), then we have the pdf*

$$
f(x) = \frac{\lambda^\beta x^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)}
$$

*So the mean of gamma distribution is:*

$$
\begin{aligned}
E[X] &= \int_0^\infty x \cdot \frac{x^{\beta-1}\lambda^\beta e^{-\lambda x}}{\Gamma(\beta)} dx = \int_0^\infty x \cdot \frac{(\lambda x)^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)} d\lambda x \\
&= \frac{1}{\lambda}\int_0^\infty \frac{(\lambda x)^{(\beta+1)-1} e^{-\lambda x} d\lambda x}{\Gamma(\beta)} = \frac{1}{\lambda\Gamma(\beta)}\Gamma(\beta+1) \\
&= \frac{\beta\Gamma(\beta)}{\lambda\Gamma(\beta)} = \frac{\beta}{\lambda}
\end{aligned}
$$

*To compute the variance, we first compute the second moment for gamma distribution:*

$$
\begin{aligned}
E[X^2] &= \int_0^\infty x^2 \cdot \frac{x^{\beta-1}\lambda^\beta e^{-\lambda x}}{\Gamma(\beta)} dx = \int_0^\infty x^2 \cdot \frac{(\lambda x)^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)} d\lambda x \\
&= \frac{1}{\lambda^2}\int_0^\infty \frac{(\lambda x)^{(\beta+2)-1} e^{-\lambda x}}{\Gamma(\beta)} d\lambda x = \frac{\Gamma(\beta+2)}{\lambda^2\Gamma(\beta)} \\
&= \frac{\beta(\beta+1)\Gamma(\beta)}{\lambda^2\Gamma(\beta)} \\
&= \frac{\beta^2+\beta}{\lambda^2}
\end{aligned}
$$

*So we now can compute the variance of gamma distribution as:*

$$
\begin{aligned}
Var[X] = E[X^2] - \Big(E[X]\Big)^2 &= \frac{\beta^2+\beta}{\lambda^2} - (\frac{\beta}{\lambda})^2 \\
&= \frac{\beta^2+\beta-\beta^2}{\lambda^2} = \frac{\beta}{\lambda^2}
\end{aligned}
$$

*(iii) If $X \sim$ lognormal$(\mu, \sigma^2)$, we have $Y = \log(X) \sim N(\mu, \sigma^2)$., which also says that $X = e^Y$. Denote $M_Y(t) = E[e^{tY}]$ as the moment generating function of $Y$, since $Y \sim N(\mu, \sigma^2)$, we have:*

$$
M_Y(t) = E[e^{tY}] = e^{\mu t + \frac{1}{2}\sigma^2 t^2}
$$

*Thus we can compute the mean of $X$ as:*

$$
E[X] = E[e^Y] = M_Y(1) = e^{\mu + \frac{1}{2}\sigma^2}
$$

*We can also compute the second moment of $X$ as:*

$$E[X^2] = E[e^{2Y}] = M_Y(2) = e^{2\mu+\frac{1}{2}\sigma^2\cdot 4} = e^{2\mu+2\sigma^2}$$

*So the variance of log normal distribution is:*

$$\begin{aligned} Var[X] &= E[X^2] - \Big(E[X]\Big)^2 = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} \\ &= e^{2\mu+\sigma^2}(e^{\sigma^2} - 1) \end{aligned}$$

*(iv) If $X \sim$ log logistic$(\alpha, \lambda)$, we have $Y = \log(X) \sim$ logistic$(\mu, \sigma)$, where $\alpha = \frac{1}{\sigma}$ and $\lambda = e^{-\frac{\mu}{\sigma}}$.*

*We denote by $M_Y(t) = E[e^{tY}]$ the moment generating function of $Y$, thus*

$$M_Y(t) = e^{\mu t} B(1 - \sigma t, 1 + \sigma t) \text{ for } \sigma t \in (-1, 1)$$

*Here $B(a, b)$ is the beta function defined as*

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$$

*So we can compute the mean of log logistic distribution as:*

$$\begin{aligned} E[X] &= E[e^Y] = M_Y(1) = e^{\mu} B(1 - \sigma, 1 + \sigma) \text{ for } \sigma \in (0, 1) \\ &= \frac{1}{\lambda^{\frac{1}{\alpha}}} B(1 - \frac{1}{\alpha}, 1 + \frac{1}{\alpha}) \text{ for } \alpha > 1 \end{aligned}$$

*We comment without proof that*

$$B(1 - t, 1 + t) = \frac{\pi t}{\sin(\pi t)} = \pi t \csc(\pi t)$$

*We also solved from $\alpha = \frac{1}{\sigma}$ and $\lambda = e^{-\frac{\mu}{\sigma}}$ that:*

$$\lambda = e^{-\alpha\mu} \Longrightarrow e^{\mu} = \frac{1}{\lambda^{\frac{1}{\alpha}}}$$

*Plug this result into the computation above, we got*

$$E[X] = \frac{1}{\lambda^{\frac{1}{\alpha}}} \cdot \frac{\pi \csc(\frac{\pi}{\alpha})}{\alpha} = \frac{\pi \csc(\pi/\alpha)}{\alpha\lambda^{\frac{1}{\alpha}}} \text{ for } \alpha > 1$$

*Now we compute the second moment:*

$$\begin{aligned} E[X^2] &= E[e^{2Y}] = M_Y(2) = e^{2\mu} B(1 - 2\sigma, 1 + 2\sigma) \text{ for } \sigma \in (0, \frac{1}{2}) \\ &= \frac{1}{\lambda^{\frac{2}{\alpha}}} B(1 - \frac{2}{\alpha}, 1 + \frac{2}{\alpha}) \text{ for } \alpha > 2 \\ &= \frac{2\pi \csc(2\pi/\alpha)}{\alpha\lambda^{\frac{2}{\alpha}}} \text{ for } \alpha > 2 \end{aligned}$$

*So the variance of log logistic distribution is:*

$$Var[X] = E[X^2] - \Big(E[X]\Big)^2 = \frac{2\pi\csc(2\pi/\alpha)}{\alpha\lambda^{\frac{2}{\alpha}}} - \Big(\frac{\pi\csc(\pi/\alpha)}{\alpha\lambda^{\frac{1}{\alpha}}}\Big)^2 \ \textit{for } \alpha > 2$$

*For part (b):*

*(i): If $X \sim \exp(\lambda)$, we have its pdf and cdf:*

$$f_X(x) = \lambda e^{-\lambda x}$$
$$F_X(x) = 1 - e^{-\lambda x}$$

*Now for $Y = \log(X)$, first consider the cdf:*

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(\log X \le y) = P(X \le e^y) = F_X(e^y) \\ &= 1 - e^{-\lambda e^y} \end{aligned}$$

*So the pdf can be computed as:*

$$f_Y(y) = \frac{d}{dy}F_Y(y) = -e^{-\lambda e^y} \cdot (-\lambda e^y) = \lambda e^{y - \lambda e^y} \ \textit{for } y \in (-\infty, +\infty)$$

*(ii): If $X \sim Weibull(\alpha, \lambda)$, we know the pdf of $X$ is:*

$$f_X(x) = \alpha\lambda x^{\alpha-1}e^{-\lambda x^\alpha}$$

*Let $Y = \log(X)$, then the cdf of $Y$ is:*

$$F_Y(y) = P(Y \le y) = P(\log X \le y) = P(X \le e^y)$$

*Thus the pdf of $Y$ can be computed as:*

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}P(X \le e^y) = \frac{d}{dy}F_X(e^y) \\ &= f_X(e^y) \cdot e^y = \alpha\lambda(e^y)^{\alpha-1} \cdot e^{-\lambda(e^y)^\alpha} \cdot e^y \\ &= \alpha\lambda e^{\alpha y} \cdot e^{-\lambda e^{\alpha y}} \end{aligned}$$

*(iii):When $X$ follows lognormal distribution, it means $Y = \log(X)$ follows normal distribution, so the pdf is:*

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

*(iv): When $X$ follows loglogistic distribution, $Y = \log(X)$ follows logistic distribution, so the pdf is*

$$f_Y(y) = \frac{e^{\frac{y-\mu}{\sigma}}}{\sigma\Big[1 + e^{\frac{y-\mu}{\sigma}}\Big]^2} \ \textit{for } y \in (-\infty, +\infty)$$

*For part (c):*

*(i) Exercise 2.3:*

*For part (a), the survival function for log-logistic distribution is*

$$S(x) = \frac{1}{1+\lambda x^{\alpha}} = \frac{1}{1+0.01x^{1.5}}$$

*So the* $50, 100,$ *and* $150$ *day survival probabilities for kidney transplantation in patiens are:*

$$\begin{aligned} S(50) &= \frac{1}{1+0.01\cdot 50^{1.5}} = 0.22 \\ S(100) &= \frac{1}{1+0.01\cdot 100^{1.5}} = 0.09 \\ S(150) &= \frac{1}{1+0.01\cdot 150^{1.5}} = 0.05 \end{aligned}$$

*For part (b):*

*At the median time to death* $x_m$ *following a kidney transplant, we have*

$$S(x_m) = \frac{1}{1+0.01x_m^{1.5}} = 0.5$$

*Solve the above equation:*

$$\begin{aligned} &1+0.01x_m^{1.5} = 2 \Longrightarrow 0.01x_m^{1.5} = 1 \\ &\Longrightarrow x_m^{1.5} = 100 \\ &\Longrightarrow x_m = 100^{\frac{2}{3}} \simeq 21.54 \end{aligned}$$

*For part (c):*

*With log logistic distribution, we have a hazard rate as:*

$$h(x) = \frac{\alpha x^{\alpha-1}\lambda}{1+\lambda x^{\alpha}} = \frac{1.5\cdot 0.01x^{0.5}}{1+0.01x^{1.5}} = \frac{0.015x^{0.5}}{1+0.01x^{1.5}}$$

*Consider the derivative of* $h(x)$*:*

$$\begin{aligned} h'(x) &= \frac{0.015\cdot\frac{1}{2\sqrt{x}}(1+0.01x^{1.5}) - 0.015x^{0.5}\cdot 0.01\cdot 1.5\sqrt{x}}{(1+0.01x^{1.5})^2} \\ &= \frac{\frac{0.015}{2\sqrt{x}} + \frac{0.00015}{2}x - (0.015)^2x}{(1+0.01x^{1.5})^2} = \frac{\left(\frac{0.00015}{2} - (0.015)^2\right)x + \frac{0.015}{2\sqrt{x}}}{(1+0.01x^{1.5})^2} \\ &= \frac{-0.00015x^{\frac{3}{2}} + 0.0075}{\sqrt{x}(1+0.01x^{1.5})^2} = \frac{-0.00015(x^{\frac{3}{2}} - 50)}{\sqrt{x}(1+0.01x^{1.5})^2} \end{aligned}$$

*From the expression above, we see that* $h'(x) > 0$ *when* $x^{\frac{3}{2}} < 50$*, and* $h(x)$ *is increasing. We also have* $h'(x) < 0$ *when* $x^{\frac{3}{2}} > 50$*, and* $h(x)$ *is decreasing. The critical point is when*

$$x^{\frac{3}{2}} = 50 \Longrightarrow x = 50^{\frac{2}{3}} \simeq 13.57$$

*The hazard rate changes from increasing to decreasing at* $x = 50^{\frac{2}{3}} = 13.57$.

*For part (d):*

*The mean time to death is:*

$$E[X] = \frac{\pi\csc(\pi/\alpha)}{\alpha\lambda^{1/\alpha}} = \frac{\pi\csc(\pi/1.5)}{1.5\cdot(0.01)^{1/1.5}} \simeq 52.10$$

*(ii) Exercise 2.5:*

*For part (a):*

*Denote time to death by* $X$*, then* $X \sim$ *lognormal*$(\mu, \sigma)$*, and* $Y = \log(X) \sim N(\mu, \sigma^2)$.

*The mean time to death is:*

$$E[X] = \exp(\mu + 0.5\sigma^2) = \exp(3.177 + 0.5\times 2.084^2) = 210.30$$

*The median time to death* $x_m$ *is computed as following:*

$$P(X > x_m) = P(Y > \log(x_m)) = P(\frac{Y-\mu}{\sigma} > \frac{\log(x_m)-\mu}{\sigma}) = P(Z > \frac{\log(x_m)-\mu}{\sigma}) = \frac{1}{2}$$

*Here* $Z$ *denote the standard normal random variable.*

*We solve the equation:*

$$\begin{aligned}
&\frac{\log(x_m)-\mu}{\sigma} = \Phi(1-0.5) = \Phi(0.5) = 0\\
&\Longrightarrow \log(x_m) = \mu = 3.177\\
&\Longrightarrow x_m = \exp(3.177) = 23.97
\end{aligned}$$

*So the median time to death is* 23.97.

*For part (b):*

*We have the survival function as:*

$$S(x) = 1 - \Phi(\frac{\log(x)-\mu}{\sigma}) = 1 - \Phi(\frac{\log(x)-3.177}{2.084})$$

*So the probability an individual survives* 100, 200 *and* 300 *days following a transplant are:*

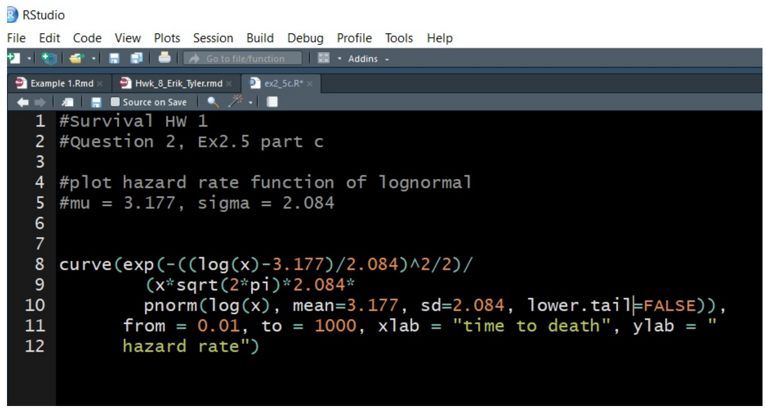
$$\begin{aligned}
S(100) &= 1 - \Phi(\frac{\log(100)-3.177}{2.084}) = 0.25\\
S(200) &= 1 - \Phi(\frac{\log(200)-3.177}{2.084}) = 0.15\\
S(300) &= 1 - \Phi(\frac{\log(300)-3.177}{2.084}) = 0.11
\end{aligned}$$

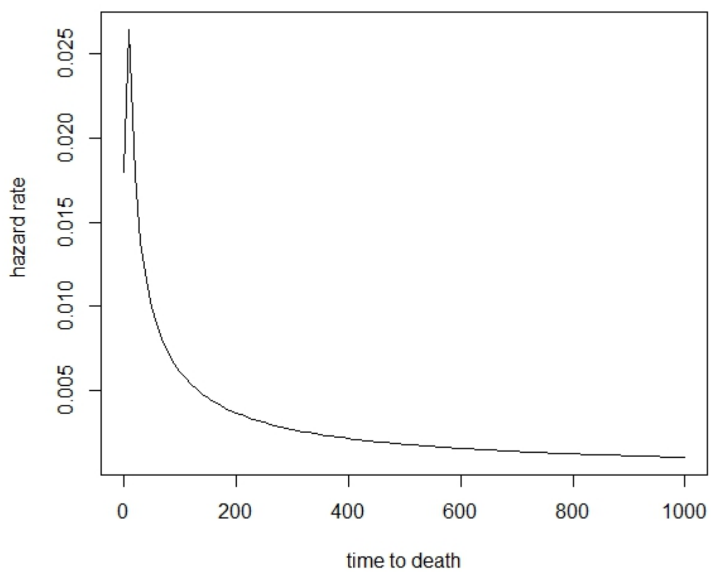
*For part (c):*

*The hazard rate of* $X$ *is:*

$$h(x) = \frac{f(x)}{S(x)} = \frac{\exp\Big[-\frac{1}{2\sigma^2}(\log(x)-\mu)^2\Big]}{x(2\pi)^{\frac{1}{2}}\sigma\Big(1-\Phi\Big[\frac{\log(x)-\mu}{\sigma}\Big]\Big)}$$

*We have the following plot:*




*We see that given $\mu = 3.177, \sigma = 2.084$, the hazard rate of log normal distribution first increase and then decrease over the time.*

*(iii) Exercise 2.7:*

*For part (a):*

*The survival function of gamma distribution is:*

$$S(x) = 1 - I(\lambda x, \beta) = 1 - \frac{\int_0^{\lambda x} u^{\beta-1} \exp(-u) du}{\Gamma(\beta)}$$
$$= 1 - \frac{\int_0^{0.2x} u^2 \exp(-u) du}{\Gamma(3)}$$

*Notice that we have:*

$$\begin{aligned}\int u^2 e^{-u} du &= -\int u^2 de^{-u} = -u^2 e^{-u} + 2\int u e^{-u} du = -u^2 e^{-u} - 2\int u de^{-u} \\ &= -u^2 e^{-u} - 2ue^{-u} + 2\int e^{-u} du = -u^2 e^{-u} - 2ue^{-u} - 2e^{-u} + constant \\ &= -e^{-u}(u^2 + 2u + 2) + constant\end{aligned}$$

*So we can explicitly compute our survival function as:*

$$\begin{aligned}S(x) &= 1 - \frac{-e^{-u}(u^2 + 2u + 2)\Big|_0^{0.2x}}{2} \\ &= 1 - \frac{-e^{-0.2x}(0.04x^2 + 0.4x + 2) + 2}{2} \\ &= \frac{1}{2} e^{-0.2x}(0.04x^2 + 0.4x + 2)\end{aligned}$$

*So the probability that a rate will survive beyond age* 18 *months is:*

$$S(18) = \frac{1}{2} e^{-0.2\times 18}(0.04 \times 18^2 + 0.4 \times 18 + 2) = 0.30$$

*part (b):*

*The probability that a rat will die in its first year of life (or not survive beyond 12 months) is:*

$$1 - S(12) = 1 - \frac{1}{2} e^{-0.2\times 12}(0.04 \times 12^2 + 0.4 \times 12 + 2) = 0.43$$

*part (c):*

*The mean lifetime for this species of rats is:*

$$E[X] = \frac{\beta}{\lambda} = \frac{3}{0.2} = 15$$

*So the mean life time is* 15 *months.*

Question #3:

**Solution 3.** *To answer textbook qustion #*3.3*:*

*For part (a):*

*a rat who had a palpable tumor at the first examination at 6 weeks after incubation with DMBA. So we know the palpable tumor happenes before at 6 weeks however the exact time we do not know. So this is left censoring.*

*For part (b):*

*For a rat that survived the study without having any tumors, if it ever develops tumors at all, it will be sometime after the end of study, and the exact time we do not know. So this is right censored.*

*For part (c):*

*For a rat which did not have a tumor at week 12 but which had a tumor at week 13 after inturbation with DMBA, it happens between two time point but the exact time we do not know, so this is an interval censoring case.*

*For part (d):*

*A rat died without tumor present and death was unrelated to the occurrence of cancer, so if the rat had lived and developed cancer, it will be sometime after the death time and the exact time of which we do not know. Thus this is a right censoring case. Since the study ending time is different than planned, we consider this as the generalized type I censoring. Since death time is independent from the time of developing tumor, this is also a random censoring.*

*For part (e):*

*The time is measured in days. Since 1 week is 7 days, so* $C_l = 6 \times 7 = 42$*. For the rat in part (a), the likelihood is then:*

$$L_a = 1 - S(C_l) = 1 - S(42)$$

*For the rat in part (b),* $C_r = 7 \times (6 + 14) = 140$*, so the likelihood is then:*

$$L_b = S(C_r) = S(140)$$

*For the rat in part (c),* $C_l = 7 \times (6 + 12) = 126$ *and* $C_r = 7 \times (6 + 13) = 133$

$$L_c = S(126) - S(133)$$

*and for the rat in part (d),* $C_r = 37$ *the likelihood is:*

$$L_d = S(37)$$

*So the total likelihood for this portion of the study is:*

$$L \propto L_a * L_b * L_c * L_d = (1 - S(42))S(140)\Big(S(126) - S(133)\Big)S(37)$$

*To answer textbook question #3.6:*

*For part (a):*

*for exponentil distribution with parameter* $\lambda$*, we know that*

$$S(x) = e^{-\lambda x}$$
$$h(x) = \lambda$$

*For relapse case, we have two different types:*

*Patient* $1-6$ *experienced the relapse event, so their likelihood is* $f(t_i), i = 1, 2, 3, 4, 5, 6$*. Patient* $7-10$ *were free of relapse by the end of study, so it is the right censoring case, so their likelihood is* $C_r(t_i)$*. Hence the likelihood for relpase is:*

$$\begin{aligned} L &\propto f(5) \times f(8) \times f(12) \times f(24) \times f(32) \times f(17) \times S(16) \times S(17) \times S(19) \times S(30) \\ &= \lambda^6 e^{-\lambda(5+8+12+24+32+17)} \times e^{-\lambda(16+17+19+30)} \\ &= \lambda^6 e^{-\lambda(98+82)} = \lambda^6 e^{-180\lambda} \end{aligned}$$

*For part (b):*

*For time to death in relapse case, patient* $1, 2, 3, 5$ *experienced death event and patient* $4, 6, 7-10$ *are right censored.*

*So the likelihood is:*

$$
\begin{aligned}
L &\propto f(11) \times f(12) \times f(15) \times f(45) \times S(33) \times S(28) \times S(16) \times S(17) \times S(19) \times S(30) \\
&= \alpha^4 \theta^4 (11 \times 12 \times 15 \times 45)^{\alpha-1} \exp\Big( -\theta(11^\alpha + 12^\alpha + 15^\alpha + 45^\alpha)\Big) \\
&\qquad \times \exp\Big[ -\theta\Big(33^\alpha + 28^\alpha + 16^\alpha + 17^\alpha + 19^\alpha + 30^\alpha\Big)\Big] \\
&= \alpha^4 \theta^4 \times 89,100^{\alpha-1} \exp\Big[ -\theta\Big(11^\alpha + 12^\alpha + 15^\alpha + 45^\alpha + 33^\alpha + 28^\alpha + 16^\alpha + 17^\alpha + 19^\alpha + 30^\alpha\Big)\Big]
\end{aligned}
$$

*For part (c):*

*Since we were only allowed to observe a patients death time if the patient relapsed, so this is a left truncated situation.*

*After the truncation we only use data from patient* $1 - 6$.

*So the likelihood is:*

$$
\begin{aligned}
L(\theta, \alpha) &\propto \frac{f(11)}{S(5)} \cdot \frac{f(12)}{S(8)} \cdot \frac{f(15)}{S(12)} \cdot \frac{S(33)}{S(24)} \cdot \frac{f(45)}{S(32)} \cdot \frac{S(28)}{S(17)} \\
&= \frac{\alpha^4 \theta^4 (11 \times 12 \times 15 \times 45)^{\alpha-1} \exp\Big( -\theta(11^\alpha + 12^\alpha + 15^\alpha + 45^\alpha)\Big)}{\exp\Big[ -\theta\Big(5^\alpha + 8^\alpha + 12^\alpha + 32^\alpha\Big)\Big]} \cdot \frac{\exp\Big[ -\theta(33^\alpha + 28^\alpha)\Big]}{\exp\Big[ -\theta\Big(24^\alpha + 17^\alpha\Big)\Big]} \\
&= \frac{\alpha^4 \theta^4 (89,100)^{\alpha-1} \exp\Big( -\theta(11^\alpha + 15^\alpha + 28^\alpha + 33^\alpha + 45^\alpha)\Big)}{\exp\Big[ -\theta\Big(5^\alpha + 8^\alpha + 17^\alpha + 24^\alpha + 32^\alpha\Big)\Big]}
\end{aligned}
$$

*We conditioned on the left truncation situation that is why we have those survival functions on the denominator, and also we did not use those subjects that did not go through relapse.*

Question #4:

**Solution 4.** *For part (a):*

*We construct the data lay out for each group:*

*For the group of AZT + zalcitabine (ddC):*

| $t_i$ | $e_i$ | $c_i$ | $n_i$ |
|---|---|---|---|
| 0 | 0 | 1 | 17 |
| 6 | 1 | 0 | 16 |
| 11 | 1 | 0 | 15 |
| 12 | 1 | 0 | 14 |
| 32 | 1 | 0 | 13 |
| 35 | 1 | 1 | 12 |
| 39 | 1 | 0 | 10 |
| 45 | 1 | 0 | 9 |
| 49 | 1 | 0 | 8 |
| 75 | 1 | 0 | 7 |
| 80 | 1 | 0 | 6 |
| 84 | 1 | 0 | 5 |
| 85 | 1 | 0 | 4 |
| 87 | 1 | 0 | 3 |
| 102 | 1 | 1 | 2 |
| *Totals* | 14 | 3 | \| |

*For the group of AZT + (ddC) + saquinivir:*

| $t_i$ | $e_i$ | $c_i$ | $n_i$ |
|---|---|---|---|
| 0 | 0 | 0 | 17 |
| 2 | 1 | 0 | 17 |
| 3 | 1 | 0 | 16 |
| 4 | 1 | 0 | 15 |
| 12 | 1 | 0 | 14 |
| 22 | 1 | 0 | 13 |
| 48 | 1 | 2 | 12 |
| 80 | 1 | 0 | 9 |
| 85 | 1 | 0 | 8 |
| 90 | 1 | 1 | 7 |
| 160 | 1 | 0 | 5 |
| 171 | 1 | 0 | 4 |
| 180 | 1 | 1 | 3 |
| 238 | 1 | 0 | 1 |
| *Totals* | 13 | 4 | \| |

*Now we compute the Kaplan-Meier estimates for each group:*

*For the AZT + zalcitabine (ddC) group:*

$$\hat{C}(0) = \frac{17}{17} = 1, \hat{S}(0) = 1$$
$$\hat{C}(6) = \frac{15}{16}, \hat{S}(6) = 1 \cdot \frac{15}{16} = 0.9375$$
$$\hat{C}(11) = \frac{14}{15}, \hat{S}(11) = \hat{S(6)} \cdot \frac{14}{15} = 0.875$$
$$\hat{C}(12) = \frac{13}{14}, \hat{S}(12) = \hat{S}(11) \cdot \frac{13}{14} = 0.8125$$
$$\hat{C}(32) = \frac{12}{13}, \hat{S}(32) = \hat{S}(12) \cdot \frac{12}{13} = 0.75$$
$$\hat{C}(35) = \frac{11}{12}, \hat{S}(35) = \hat{S}(32) \cdot \frac{11}{12} = 0.6875$$
$$\hat{C}(39) = \frac{9}{10}, \hat{S}(39) = \hat{S}(35) \cdot \frac{9}{10} = 0.61875$$
$$\hat{C}(45) = \frac{8}{9}, \hat{S}(45) = \hat{S}(39) \cdot \frac{8}{9} = 0.55$$
$$\hat{C}(49) = \frac{7}{8}, \hat{S}(49) = \hat{S}(45) \cdot \frac{7}{8} = 0.48125$$
$$\hat{C}(75) = \frac{6}{7}, \hat{S}(75) = \hat{S}(49) \cdot \frac{6}{7} = 0.4125$$
$$\hat{C}(80) = \frac{5}{6}, \hat{S}(80) = \hat{S}(75) \cdot \frac{5}{6} = 0.34375$$
$$\hat{C}(84) = \frac{4}{5}, \hat{S}(84) = \hat{S}(80) \cdot \frac{4}{5} = 0.275$$
$$\hat{C}(85) = \frac{3}{4}, \hat{S}(85) = \hat{S}(84) \cdot \frac{3}{4} = 0.20625$$
$$\hat{C}(87) = \frac{2}{3}, \hat{S}(87) = \hat{S}(85) \cdot \frac{2}{3} = 0.1375$$
$$\hat{C}(102) = \frac{1}{2}, \hat{S}(102) = \hat{S}(87) \cdot \frac{1}{2} = 0.06875$$

*We incorporate this information into the previous data layout to get:*

| $t_i$ | $e_i$ | $c_i$ | $n_i$ | $\hat{S}(t_{j-1})$ | $\hat{C}(t_j)$ | $\hat{S}(t_j)$ |
|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 17 | - | $\frac{17}{17}$ | 1 |
| 6 | 1 | 0 | 16 | 1 | $\frac{15}{16}$ | 0.9375 |
| 11 | 1 | 0 | 15 | 0.9375 | $\frac{14}{15}$ | 0.875 |
| 12 | 1 | 0 | 14 | 0.875 | $\frac{13}{14}$ | 0.8125 |
| 32 | 1 | 0 | 13 | 0.8125 | $\frac{12}{13}$ | 0.75 |
| 35 | 1 | 1 | 12 | 0.75 | $\frac{11}{12}$ | 0.6875 |
| 39 | 1 | 0 | 10 | 0.6875 | $\frac{9}{10}$ | 0.61875 |
| 45 | 1 | 0 | 9 | 0.61875 | $\frac{8}{9}$ | 0.55 |
| 49 | 1 | 0 | 8 | 0.55 | $\frac{7}{8}$ | 0.48125 |
| 75 | 1 | 0 | 7 | 0.48125 | $\frac{6}{7}$ | 0.4125 |
| 80 | 1 | 0 | 6 | 0.4125 | $\frac{5}{6}$ | 0.34375 |
| 84 | 1 | 0 | 5 | 0.34375 | $\frac{4}{5}$ | 0.275 |
| 85 | 1 | 0 | 4 | 0.275 | $\frac{3}{4}$ | 0.20625 |
| 87 | 1 | 0 | 3 | 0.20625 | $\frac{2}{3}$ | 0.1375 |
| 102 | 1 | 1 | 2 | 0.1375 | $\frac{1}{2}$ | 0.06875 |
| *Totals* | 14 | 3 | \| | | | |

*Similarly for the AZT + zalcitabine (ddC) + saquinivir group:*

$$\hat{C}(0) = \frac{17}{17} = 1, \hat{S}(0) = 1$$
$$\hat{C}(2) = \frac{16}{17}, \hat{S}(2) = 1 \cdot \frac{16}{17} = 0.94$$
$$\hat{C}(3) = \frac{15}{16}, \hat{S}(3) = \hat{S}(2) \cdot \frac{15}{16} = 0.88$$
$$\hat{C}(4) = \frac{14}{15}, \hat{S}(4) = \hat{S}(3) \cdot \frac{14}{15} = 0.82$$
$$\hat{C}(12) = \frac{13}{14}, \hat{S}(12) = \hat{S}(4) \cdot \frac{13}{14} = 0.765$$
$$\hat{C}(22) = \frac{12}{13}, \hat{S}(22) = \hat{S}(12) \cdot \frac{12}{13} = 0.71$$
$$\hat{C}(48) = \frac{11}{12}, \hat{S}(48) = \hat{S}(22) \cdot \frac{11}{12} = 0.65$$
$$\hat{C}(80) = \frac{8}{9}, \hat{S}(80) = \hat{S}(48) \cdot \frac{8}{9} = 0.58$$
$$\hat{C}(85) = \frac{7}{8}, \hat{S}(85) = \hat{S}(80) \cdot \frac{7}{8} = 0.503$$
$$\hat{C}(90) = \frac{6}{7}, \hat{S}(90) = \hat{S}(85) \cdot \frac{6}{7} = 0.43$$

$$\hat{C}(160) = \frac{4}{5}, \hat{S}(160) = \hat{S}(90) \cdot \frac{4}{5} = 0.345$$
$$\hat{C}(171) = \frac{3}{4}, \hat{S}(171) = \hat{S}(160) \cdot \frac{3}{4} = 0.259$$
$$\hat{C}(180) = \frac{2}{3}, \hat{S}(180) = \hat{S}(171) \cdot \frac{2}{3} = 0.17$$
$$\hat{C}(238) = 0, \hat{S}(238) = 0$$

*Incorporate this informatino into the data layout above, we got: For the group of AZT + (ddC) + saquinivir:*

| $t_i$ | $e_i$ | $c_i$ | $n_i$ | $\hat{S}(t_{j-1})$ | $\hat{C}(t_j)$ | $\hat{S}(t_j)$ |
|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 17 | - | *1* | *1* |
| 2 | 1 | 0 | 17 | 1 | $\frac{16}{17}$ | 0.94 |
| 3 | 1 | 0 | 16 | 0.94 | $\frac{15}{16}$ | 0.88 |
| 4 | 1 | 0 | 15 | 0.88 | $\frac{14}{15}$ | 0.82 |
| 12 | 1 | 0 | 14 | 0.82 | $\frac{13}{14}$ | 0.765 |
| 22 | 1 | 0 | 13 | 0.765 | $\frac{12}{13}$ | 0.71 |
| 48 | 1 | 2 | 12 | 0.71 | $\frac{11}{12}$ | 0.65 |
| 80 | 1 | 0 | 9 | 0.65 | $\frac{8}{9}$ | 0.58 |
| 85 | 1 | 0 | 8 | 0.58 | $\frac{7}{8}$ | 0.503 |
| 90 | 1 | 1 | 7 | 0.503 | $\frac{6}{7}$ | 0.43 |
| 160 | 1 | 0 | 5 | 0.43 | $\frac{4}{5}$ | 0.345 |
| 171 | 1 | 0 | 4 | 0.345 | $\frac{3}{4}$ | 0.259 |
| 180 | 1 | 1 | 3 | 0.259 | $\frac{2}{3}$ | 0.17 |
| 238 | 1 | 0 | 1 | 0.17 | 0 | 0 |
| *Totals* | 13 | 4 | \| | | | |

*The complete KM estimates for each group and for any t betwee the starting time and the largest time is a step function whose values are defined by the values of $\hat{S}(t_j)$ in the two tables above.*

*Based on the two tables above:*

*The median survival time(point estimate) for group AZT + zalcitabine (ddC) is $t = 49$ since $\hat{S}(45) = 0.55$ and $\hat{S}(49) = 0.48125$.*

*The median survival time(point estimate) for group AZT + (ddC) + saquinivir is $t = 90$ since $\hat{S}(85) = 0.503$ and $\hat{S}(90) = 0.43$.*

*I would not feel comfortable reporting the mean survival time for group AZT + ddc because there is censoring time greater than the largest event time, so the estimated mean survival time will be biased downward.*

*However we can report the mean survival timefor the second group AZT + ddc + saquinivir.*

*For part (b):*

*We run the following SAS code to verify that we got the correct Kaplan-Meier estimates in part (a):*

```
dm 'log; clear; output; clear;';


proc format;
    value status 0 = "censored"
                 1 = "event";
    value group  1 = "AZT + ddc"
                 2 = "AZT + ddc + saquinivir";
;

data Q4;
    input group time status @@;
    format group group.
           status status.;
    datalines;
    1 4 0   1 6 1   1 11 1  1 12 1  1 32 1  1 35 1
    1 38 0  1 39 1  1 45 1  1 49 1  1 75 1  1 80 1
    1 84 1  1 85 1  1 87 1  1 102 1 1 180 0 2 2 1
    2 3 1   2 4 1   2 12 1  2 22 1  2 48 1  2 51 0
    2 56 0  2 80 1  2 85 1  2 90 1  2 94 0  2 160 1
    2 171 1 2 180 1 2 180 0 2 238 1
    ;
run;

proc print data=Q4;
run;

title "Question 4 KM Estimate";

proc lifetest data=Q4;
    time time*status(0);
    by group;
run;
```

*Then we got the Kaplan Meier estimates for group AZT + ddc as the following:*

**Question 4 KM Estimate**

The LIFETEST Procedure

group=AZT + ddc

| Product-Limit Survival Estimates | | | | | | |
|---|---|---|---|---|---|---|
| time | | Survival | Failure | Survival Standard Error | Number Failed | Number Left |
| 0.000 | | 1.0000 | 0 | 0 | 0 | 17 |
| 4.000 | * | . | . | . | 0 | 16 |
| 6.000 | | 0.9375 | 0.0625 | 0.0605 | 1 | 15 |
| 11.000 | | 0.8750 | 0.1250 | 0.0827 | 2 | 14 |
| 12.000 | | 0.8125 | 0.1875 | 0.0976 | 3 | 13 |
| 32.000 | | 0.7500 | 0.2500 | 0.1083 | 4 | 12 |
| 35.000 | | 0.6875 | 0.3125 | 0.1159 | 5 | 11 |
| 38.000 | * | . | . | . | 5 | 10 |
| 39.000 | | 0.6188 | 0.3813 | 0.1230 | 6 | 9 |
| 45.000 | | 0.5500 | 0.4500 | 0.1271 | 7 | 8 |
| 49.000 | | 0.4813 | 0.5188 | 0.1285 | 8 | 7 |
| 75.000 | | 0.4125 | 0.5875 | 0.1272 | 9 | 6 |
| 80.000 | | 0.3438 | 0.6563 | 0.1232 | 10 | 5 |
| 84.000 | | 0.2750 | 0.7250 | 0.1162 | 11 | 4 |
| 85.000 | | 0.2063 | 0.7938 | 0.1055 | 12 | 3 |
| 87.000 | | 0.1375 | 0.8625 | 0.0900 | 13 | 2 |
| 102.000 | | 0.0688 | 0.9313 | 0.0662 | 14 | 1 |
| 180.000 | * | 0.0688 | 0.9313 | . | 14 | 0 |

*We also have the Kaplan Meier estimates for group AZT + ddc + saquinivir as :*

**The LIFETEST Procedure**

**group=AZT + ddc + saquinivir**

| Product-Limit Survival Estimates | | | | | | |
|---|---|---|---|---|---|---|
| time | | Survival | Failure | Survival Standard Error | Number Failed | Number Left |
| 0.000 | | 1.0000 | 0 | 0 | 0 | 17 |
| 2.000 | | 0.9412 | 0.0588 | 0.0571 | 1 | 16 |
| 3.000 | | 0.8824 | 0.1176 | 0.0781 | 2 | 15 |
| 4.000 | | 0.8235 | 0.1765 | 0.0925 | 3 | 14 |
| 12.000 | | 0.7647 | 0.2353 | 0.1029 | 4 | 13 |
| 22.000 | | 0.7059 | 0.2941 | 0.1105 | 5 | 12 |
| 48.000 | | 0.6471 | 0.3529 | 0.1159 | 6 | 11 |
| 51.000 | * | . | . | . | 6 | 10 |
| 56.000 | * | . | . | . | 6 | 9 |
| 80.000 | | 0.5752 | 0.4248 | 0.1233 | 7 | 8 |
| 85.000 | | 0.5033 | 0.4967 | 0.1272 | 8 | 7 |
| 90.000 | | 0.4314 | 0.5686 | 0.1277 | 9 | 6 |
| 94.000 | * | . | . | . | 9 | 5 |
| 160.000 | | 0.3451 | 0.6549 | 0.1280 | 10 | 4 |
| 171.000 | | 0.2588 | 0.7412 | 0.1217 | 11 | 3 |
| 180.000 | | 0.1725 | 0.8275 | 0.1074 | 12 | 2 |
| 180.000 | * | . | . | . | 12 | 1 |
| 238.000 | | 0 | 1.0000 | . | 13 | 0 |

**Note:** The marked survival times are censored observations.

*Compare with the manually computed result in part (a), they do match.*

*For part (c):*

*We give the following SAS code:*

```
proc lifetest data=Q4;
    time time*status(0);
    by group;
run;

title "Compare H hat and H tilde";
proc lifetest data=Q4 nelson;
    time time*status(0);
    by group;
    ods output productlimitestimates=ple;
run;

data cumhazard;
    set ple;
    hatcumhaz = -log(Survival);
run;

proc print data=cumhazard;
run;

title "cuulative hazard fuction estimate comparison";
title2 "nelson-aalen estimate and product limit estimate";
proc sgpanel data=cumhazard
    (where = (censor = 0 and (group = 1 or group = 2)));
    panelby group;
    step x =time y = CumHaz/legendlabel = "N-A estimates";
    step x = time y = hatCumHaz/legendlabel = "product limit estimates";
    rowaxis label = "cumulative hazard function estimate";
run;
```

*The first proc lifetest gives the following out put by group for the cumulative hazard fuction estimate (based on Nelson Aalen):*

*For group AZT + ddc:*

group=AZT + ddc

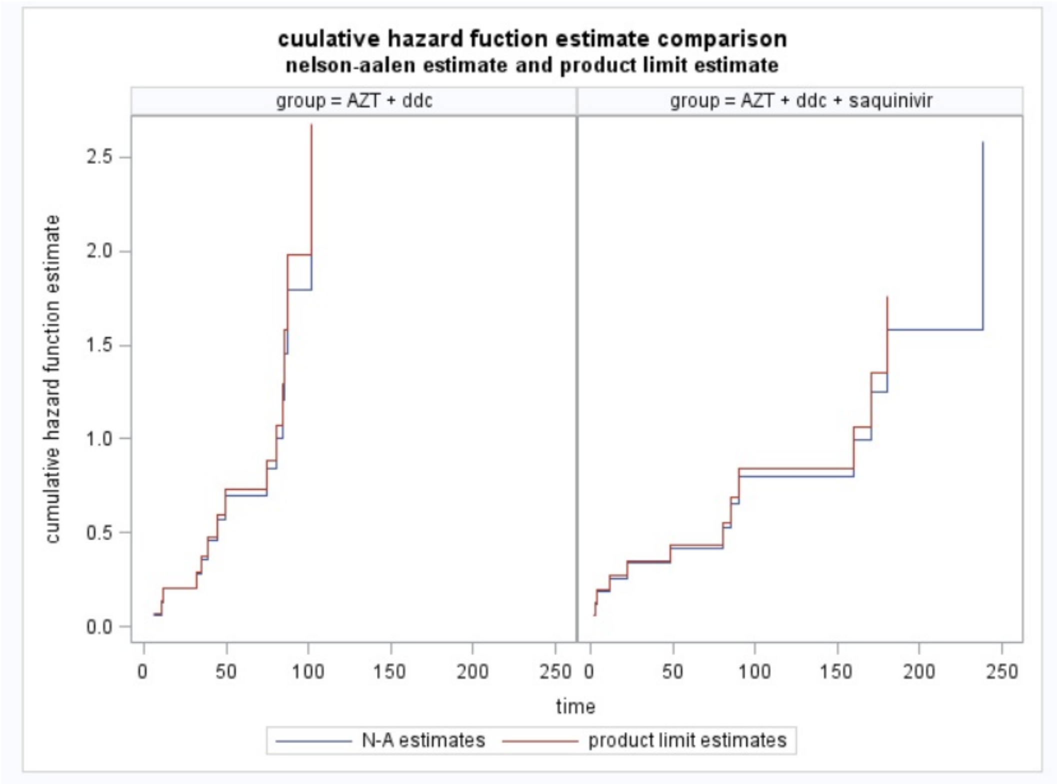
| Survival Function and Cumulative Hazard Rate | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Product-Limit | | | Nelson-Aalen | | | |
| time | | Survival | Failure | Survival Standard Error | Cumulative Hazard | Cum Haz Standard Error | Number Failed | Number Left |
| 0.000 | | 1.0000 | 0 | 0 | 0 | . | 0 | 17 |
| 4.000 | * | . | . | . | . | . | 0 | 16 |
| 6.000 | | 0.9375 | 0.0625 | 0.0605 | 0.0625 | 0.0625 | 1 | 15 |
| 11.000 | | 0.8750 | 0.1250 | 0.0827 | 0.1292 | 0.0914 | 2 | 14 |
| 12.000 | | 0.8125 | 0.1875 | 0.0976 | 0.2006 | 0.1160 | 3 | 13 |
| 32.000 | | 0.7500 | 0.2500 | 0.1083 | 0.2775 | 0.1392 | 4 | 12 |
| 35.000 | | 0.6875 | 0.3125 | 0.1159 | 0.3609 | 0.1622 | 5 | 11 |
| 38.000 | * | . | . | . | . | . | 5 | 10 |
| 39.000 | | 0.6188 | 0.3813 | 0.1230 | 0.4609 | 0.1906 | 6 | 9 |
| 45.000 | | 0.5500 | 0.4500 | 0.1271 | 0.5720 | 0.2206 | 7 | 8 |
| 49.000 | | 0.4813 | 0.5188 | 0.1285 | 0.6970 | 0.2535 | 8 | 7 |
| 75.000 | | 0.4125 | 0.5875 | 0.1272 | 0.8398 | 0.2910 | 9 | 6 |
| 80.000 | | 0.3438 | 0.6563 | 0.1232 | 1.0065 | 0.3354 | 10 | 5 |
| 84.000 | | 0.2750 | 0.7250 | 0.1162 | 1.2065 | 0.3905 | 11 | 4 |
| 85.000 | | 0.2063 | 0.7938 | 0.1055 | 1.4565 | 0.4636 | 12 | 3 |
| 87.000 | | 0.1375 | 0.8625 | 0.0900 | 1.7898 | 0.5710 | 13 | 2 |
| 102.000 | | 0.0688 | 0.9313 | 0.0662 | 2.2898 | 0.7590 | 14 | 1 |
| 180.000 | * | 0.0688 | 0.9313 | . | . | . | 14 | 0 |

*For group AZT + ddc + saquinivir:*

group=AZT + ddc + saquinivir

| Survival Function and Cumulative Hazard Rate | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Product-Limit | | | Nelson-Aalen | | | |
| time | | Survival | Failure | Survival Standard Error | Cumulative Hazard | Cum Haz Standard Error | Number Failed | Number Left |
| 0.000 | | 1.0000 | 0 | 0 | 0 | . | 0 | 17 |
| 2.000 | | 0.9412 | 0.0588 | 0.0571 | 0.0588 | 0.0588 | 1 | 16 |
| 3.000 | | 0.8824 | 0.1176 | 0.0781 | 0.1213 | 0.0858 | 2 | 15 |
| 4.000 | | 0.8235 | 0.1765 | 0.0925 | 0.1880 | 0.1087 | 3 | 14 |
| 12.000 | | 0.7647 | 0.2353 | 0.1029 | 0.2594 | 0.1300 | 4 | 13 |
| 22.000 | | 0.7059 | 0.2941 | 0.1105 | 0.3363 | 0.1511 | 5 | 12 |
| 48.000 | | 0.6471 | 0.3529 | 0.1159 | 0.4197 | 0.1726 | 6 | 11 |
| 51.000 | * | . | . | . | . | . | 6 | 10 |
| 56.000 | * | . | . | . | . | . | 6 | 9 |
| 80.000 | | 0.5752 | 0.4248 | 0.1233 | 0.5308 | 0.2052 | 7 | 8 |
| 85.000 | | 0.5033 | 0.4967 | 0.1272 | 0.6558 | 0.2403 | 8 | 7 |
| 90.000 | | 0.4314 | 0.5686 | 0.1277 | 0.7986 | 0.2796 | 9 | 6 |
| 94.000 | * | . | . | . | . | . | 9 | 5 |
| 160.000 | | 0.3451 | 0.6549 | 0.1280 | 0.9986 | 0.3437 | 10 | 4 |
| 171.000 | | 0.2588 | 0.7412 | 0.1217 | 1.2486 | 0.4250 | 11 | 3 |
| 180.000 | | 0.1725 | 0.8275 | 0.1074 | 1.5820 | 0.5402 | 12 | 2 |
| 180.000 | * | . | . | . | . | . | 12 | 1 |
| 238.000 | | 0 | 1.0000 | . | 2.5820 | 1.1366 | 13 | 0 |

*The second proc lifetest gives a plot of comparison between the cumulative hazard function estimate acquired from product limit (Kaplan Meier) and from Nelson-Aalen(N-A):*




*and as expected we see that in both groups we have $\hat{H}(t) \geq \tilde{H}(t)$, this is because we have:*

$$\exp[-\hat{H}(t)] = \hat{S}(t) = \prod_{t_i \leq t} \left[1 - \frac{d_i}{Y_i}\right] \leq \prod_{t_i \leq t} \exp\left[-\frac{d_i}{Y_i}\right] = \exp[-\sum_{t_i \leq t} \frac{d_i}{Y_i}] = \exp[-\tilde{H}(t)] = \tilde{S}(t)$$

*which leads to*

$$-\hat{H}(t) \leq -\tilde{H}(t)$$

*and hence*

$$\hat{H}(t) \geq \tilde{H}(t)$$

*For part (d):*

*We give the following SAS code to produce life-table estimate of the survival function and to plot the life table estimates of the hazard function:*

```
title "life table estimate";
proc lifetest data=Q4 method = life width = 60 plots = (H);
    time time*status(0);
    by group;
run;
```

*partiularly the width = 60 option specify that the estimate is based on interval width of* 60 *days.*

*For group AZT + ddc we have the following output, and among which the survival column gives the life table estimate for survival function based on interval width of* 60 *days.*

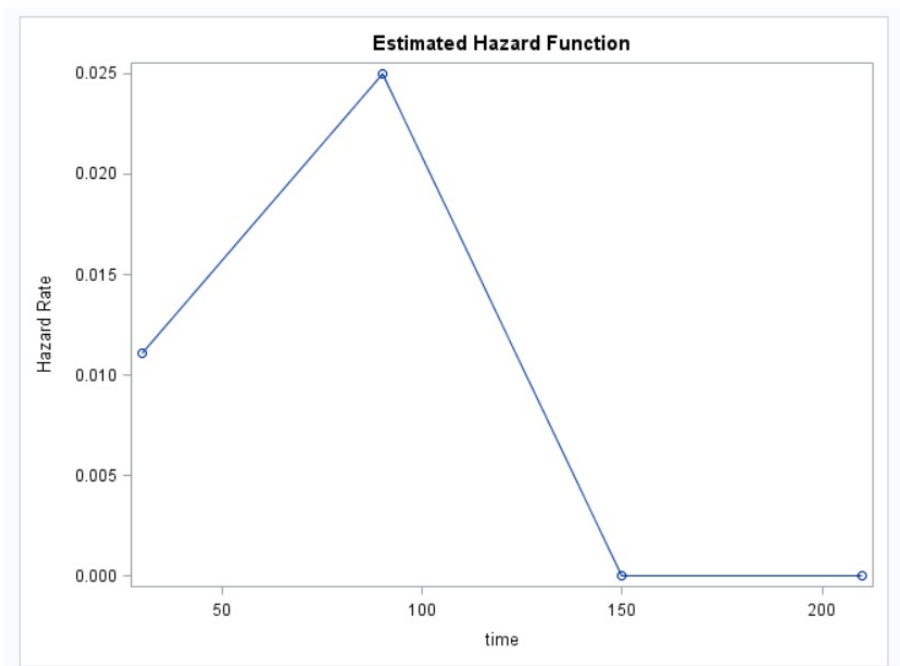
**life table estimate**

The LIFETEST Procedure

group=AZT + ddc

Life Table Survival Estimates

| Interval | | | | | | | | | | | | Evaluated at the Midpoint of the Interval | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error | Median Residual Lifetime | Median Standard Error | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
| 0 | 60 | 8 | 2 | 16.0 | 0.5000 | 0.1250 | 1.0000 | 0 | 0 | 60.0000 | 15.0000 | 0.00833 | 0.00208 | 0.011111 | 0.003704 |
| 60 | 120 | 6 | 0 | 7.0 | 0.8571 | 0.1323 | 0.5000 | 0.5000 | 0.1250 | 35.0000 | 13.2288 | 0.00714 | 0.00210 | 0.025 | 0.006751 |
| 120 | 180 | 0 | 0 | 1.0 | 0 | 0 | 0.0714 | 0.9286 | 0.0685 | . | . | 0 | . | 0 | . |
| 180 | 240 | 0 | 1 | 0.5 | 0 | 0 | 0.0714 | 0.9286 | 0.0685 | . | . | 0 | . | 0 | . |
| 240 | . | 0 | 0 | 0.0 | 0 | 0 | 0.0714 | 0.9286 | 0.0685 | . | . | . | . | . | . |




*The hazard rate reaches a peak but then decrease with respect to the time, indicating the drug resistance later on and less efficacy over long time.*

*For group AZT + ddc + saquivinir,we have the following output,*

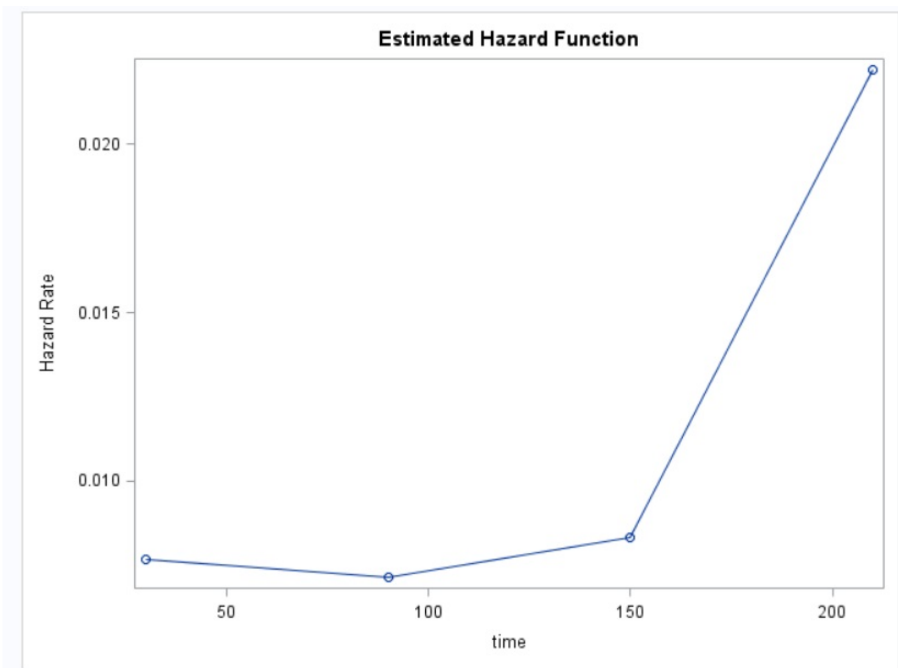
**life table estimate**

The LIFETEST Procedure

group=AZT + ddc + saquinivir

Life Table Survival Estimates

| Interval | | | | | | | | | | | | Evaluated at the Midpoint of the Interval | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error | Median Residual Lifetime | Median Standard Error | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
| 0 | 60 | 6 | 2 | 16.0 | 0.3750 | 0.1210 | 1.0000 | 0 | 0 | 94.0000 | 34.0000 | 0.00625 | 0.00202 | 0.007692 | 0.003056 |
| 60 | 120 | 3 | 1 | 8.5 | 0.3529 | 0.1639 | 0.6250 | 0.3750 | 0.1210 | 94.0909 | 39.7565 | 0.00368 | 0.00185 | 0.007143 | 0.004028 |
| 120 | 180 | 2 | 0 | 5.0 | 0.4000 | 0.2191 | 0.4044 | 0.5956 | 0.1290 | 72.5000 | 27.9508 | 0.00270 | 0.00171 | 0.008333 | 0.005705 |
| 180 | 240 | 2 | 1 | 2.5 | 0.8000 | 0.2530 | 0.2426 | 0.7574 | 0.1176 | 37.5000 | 23.7171 | 0.00324 | 0.00187 | 0.022222 | 0.011712 |
| 240 | . | 0 | 0 | 0.0 | 0 | 0 | 0.0485 | 0.9515 | 0.0657 | . | . | . | . | . | . |




*the hazard rate keep increasing later on, indicates a long term efficancy and less of a drug resistance*

*For part (e):*

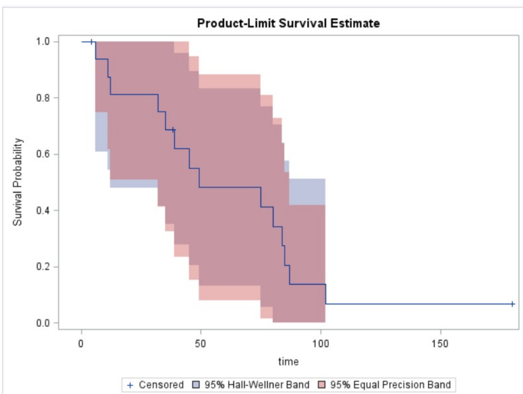
*For example, if we would like to build confidence interval under linear transformation, and sketch the confidence band of both equal probaiblity and hall wellner type, we have the following SAS code:*

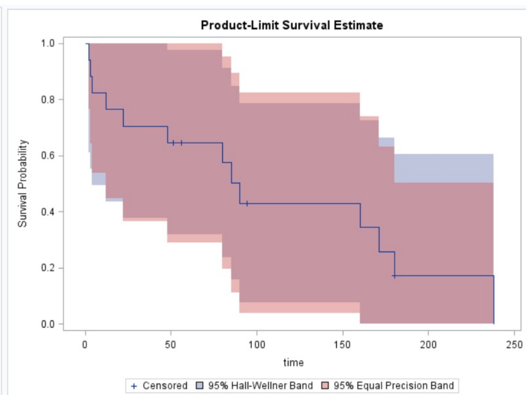
Guanlin Zhang
Spring '18

```
title "confidence bands";
title2 "linear transformation";
proc lifetest data=Q4
      outsurv = results1 conftype = linear confband = all
      plots = survival(CB = all);
     time time*status(0);
     by group;
run;

proc print data=results1;
run;
```

*We have the following plots for each group:*




*and the following table of confidence intervals*

**confidence bands**
**linear transformation**

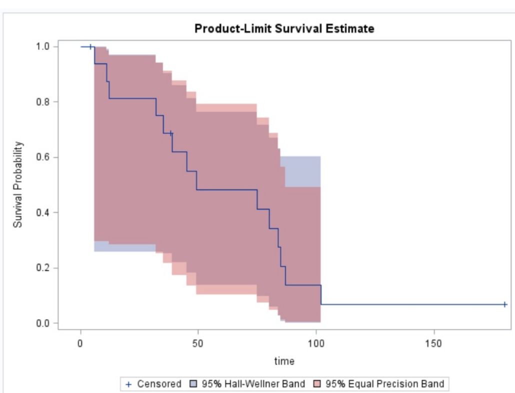
| Obs | group | time | _CENSOR_ | SURVIVAL | SDF_LCL | SDF_UCL | HW_LCL | HW_UCL | EP_LCL | EP_UCL |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AZT + ddc | 0 | . | 1.00000 | 1.00000 | 1.00000 | . | . | . | . |
| 2 | AZT + ddc | 4 | 1 | 1.00000 | . | . | . | . | . | . |
| 3 | AZT + ddc | 6 | 0 | 0.93750 | 0.81889 | 1.00000 | 0.60683 | 1.00000 | 0.74852 | 1.00000 |
| 4 | AZT + ddc | 11 | 0 | 0.87500 | 0.71295 | 1.00000 | 0.54304 | 1.00000 | 0.61680 | 1.00000 |
| 5 | AZT + ddc | 12 | 0 | 0.81250 | 0.62125 | 1.00000 | 0.47925 | 1.00000 | 0.50778 | 1.00000 |
| 6 | AZT + ddc | 32 | 0 | 0.75000 | 0.53783 | 0.96217 | 0.41547 | 1.00000 | 0.41194 | 1.00000 |
| 7 | AZT + ddc | 35 | 0 | 0.68750 | 0.46038 | 0.91462 | 0.35168 | 1.00000 | 0.32563 | 1.00000 |
| 8 | AZT + ddc | 38 | 1 | 0.68750 | . | . | . | . | . | . |
| 9 | AZT + ddc | 39 | 0 | 0.61875 | 0.37766 | 0.85984 | 0.27801 | 0.95949 | 0.23462 | 1.00000 |
| 10 | AZT + ddc | 45 | 0 | 0.55000 | 0.30087 | 0.79913 | 0.20435 | 0.89565 | 0.15306 | 0.94694 |
| 11 | AZT + ddc | 49 | 0 | 0.48125 | 0.22945 | 0.73305 | 0.13068 | 0.83182 | 0.08005 | 0.88245 |
| 12 | AZT + ddc | 75 | 0 | 0.41250 | 0.16321 | 0.66179 | 0.05702 | 0.76798 | 0.01530 | 0.80970 |
| 13 | AZT + ddc | 80 | 0 | 0.34375 | 0.10232 | 0.58518 | 0.00000 | 0.70415 | 0.00000 | 0.72842 |
| 14 | AZT + ddc | 84 | 0 | 0.27500 | 0.04734 | 0.50266 | 0.00000 | 0.64031 | 0.00000 | 0.63774 |
| 15 | AZT + ddc | 85 | 0 | 0.20625 | 0.00000 | 0.41306 | 0.00000 | 0.57648 | 0.00000 | 0.53577 |
| 16 | AZT + ddc | 87 | 0 | 0.13750 | 0.00000 | 0.31389 | 0.00000 | 0.51264 | 0.00000 | 0.41855 |
| 17 | AZT + ddc | 102 | 0 | 0.06875 | 0.00000 | 0.19858 | 0.00000 | 0.44881 | 0.00000 | 0.27562 |
| 18 | AZT + ddc | 180 | 1 | . | . | . | . | . | . | . |

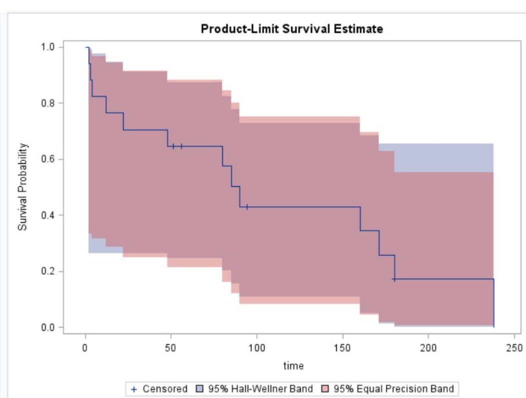
| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 19 | AZT + ddc + saquinivir | 0 | . | 1.00000 | 1.00000 | 1.00000 | . | . | . | . |
| 20 | AZT + ddc + saquinivir | 2 | 0 | 0.94118 | 0.82933 | 1.00000 | 0.61181 | 1.00000 | 0.76578 | 1.00000 |
| 21 | AZT + ddc + saquinivir | 3 | 0 | 0.88235 | 0.72920 | 1.00000 | 0.55298 | 1.00000 | 0.64218 | 1.00000 |
| 22 | AZT + ddc + saquinivir | 4 | 0 | 0.82353 | 0.64231 | 1.00000 | 0.49416 | 1.00000 | 0.53936 | 1.00000 |
| 23 | AZT + ddc + saquinivir | 12 | 0 | 0.76471 | 0.56307 | 0.96635 | 0.43533 | 1.00000 | 0.44851 | 1.00000 |
| 24 | AZT + ddc + saquinivir | 22 | 0 | 0.70588 | 0.48929 | 0.92248 | 0.37651 | 1.00000 | 0.36623 | 1.00000 |
| 25 | AZT + ddc + saquinivir | 48 | 0 | 0.64706 | 0.41989 | 0.87423 | 0.31769 | 0.97643 | 0.29083 | 1.00000 |
| 26 | AZT + ddc + saquinivir | 51 | 1 | 0.64706 | . | . | . | . | . | . |
| 27 | AZT + ddc + saquinivir | 56 | 1 | 0.64706 | . | . | . | . | . | . |
| 28 | AZT + ddc + saquinivir | 80 | 0 | 0.57516 | 0.33345 | 0.81688 | 0.23766 | 0.91267 | 0.19613 | 0.95420 |
| 29 | AZT + ddc + saquinivir | 85 | 0 | 0.50327 | 0.25406 | 0.75248 | 0.15763 | 0.84890 | 0.11247 | 0.89406 |
| 30 | AZT + ddc + saquinivir | 90 | 0 | 0.43137 | 0.18108 | 0.68167 | 0.07760 | 0.78514 | 0.03887 | 0.82387 |
| 31 | AZT + ddc + saquinivir | 94 | 1 | 0.43137 | . | . | . | . | . | . |
| 32 | AZT + ddc + saquinivir | 160 | 0 | 0.34510 | 0.09416 | 0.59604 | 0.00000 | 0.72473 | 0.00000 | 0.73860 |
| 33 | AZT + ddc + saquinivir | 171 | 0 | 0.25882 | 0.02036 | 0.49729 | 0.00000 | 0.66432 | 0.00000 | 0.63277 |
| 34 | AZT + ddc + saquinivir | 180 | 0 | 0.17255 | 0.00000 | 0.38311 | 0.00000 | 0.60390 | 0.00000 | 0.50273 |
| 35 | AZT + ddc + saquinivir | 180 | 1 | 0.17255 | . | . | . | . | . | . |
| 36 | AZT + ddc + saquinivir | 238 | 0 | 0.00000 | 0.00000 | 0.00000 | . | . | . | . |

*We can change the option of confidence type, so for log-log transformation, we have*

```
title2 "log-log transformation";
proc lifetest data=Q4
     outsurv = results2 conftype = loglog confband = all
     plots = survival(CB = all);
    time time*status(0);
    by group;
run;

proc print data=results2;
run;
```




*and*

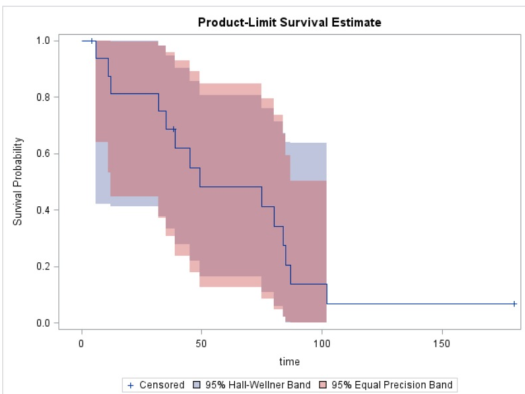
**confidence bands**
**log-log transformation**

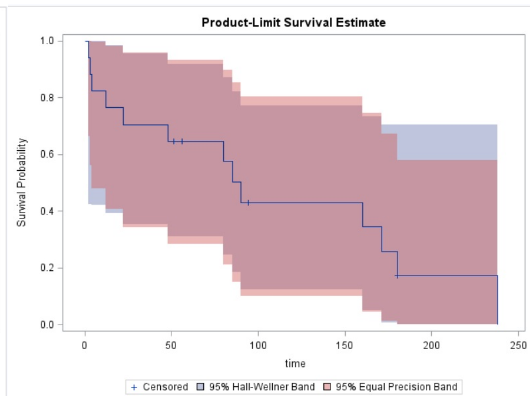
| Obs | group | time | _CENSOR_ | SURVIVAL | SDF_LCL | SDF_UCL | HW_LCL | HW_UCL | EP_LCL | EP_UCL |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AZT + ddc | 0 | . | 1.00000 | 1.00000 | 1.00000 | . | . | . | . |
| 2 | AZT + ddc | 4 | 1 | 1.00000 | . | . | . | . | . | . |
| 3 | AZT + ddc | 6 | 0 | 0.93750 | 0.63235 | 0.99095 | 0.00000 | 0.99973 | 0.23073 | 0.99716 |
| 4 | AZT + ddc | 11 | 0 | 0.87500 | 0.58598 | 0.96719 | 0.10146 | 0.99224 | 0.29611 | 0.98546 |
| 5 | AZT + ddc | 12 | 0 | 0.81250 | 0.52460 | 0.93535 | 0.22384 | 0.97161 | 0.28253 | 0.96647 |
| 6 | AZT + ddc | 32 | 0 | 0.75000 | 0.46343 | 0.89798 | 0.25767 | 0.94079 | 0.25199 | 0.94172 |
| 7 | AZT + ddc | 35 | 0 | 0.68750 | 0.40460 | 0.85628 | 0.25161 | 0.90326 | 0.21725 | 0.91214 |
| 8 | AZT + ddc | 38 | 1 | 0.68750 | . | . | . | . | . | . |
| 9 | AZT + ddc | 39 | 0 | 0.61875 | 0.33929 | 0.80799 | 0.22052 | 0.85861 | 0.17385 | 0.87659 |
| 10 | AZT + ddc | 45 | 0 | 0.55000 | 0.27933 | 0.75560 | 0.18078 | 0.81143 | 0.13544 | 0.83629 |
| 11 | AZT + ddc | 49 | 0 | 0.48125 | 0.22410 | 0.69933 | 0.13805 | 0.76328 | 0.10162 | 0.79141 |
| 12 | AZT + ddc | 75 | 0 | 0.41250 | 0.17339 | 0.63922 | 0.09600 | 0.71561 | 0.07230 | 0.74193 |
| 13 | AZT + ddc | 80 | 0 | 0.34375 | 0.12728 | 0.57512 | 0.05782 | 0.67029 | 0.04758 | 0.68767 |
| 14 | AZT + ddc | 84 | 0 | 0.27500 | 0.08617 | 0.50668 | 0.02699 | 0.63043 | 0.02770 | 0.62831 |
| 15 | AZT + ddc | 85 | 0 | 0.20625 | 0.05082 | 0.43325 | 0.00729 | 0.60268 | 0.01299 | 0.56338 |
| 16 | AZT + ddc | 87 | 0 | 0.13750 | 0.02265 | 0.35367 | 0.00039 | 0.60554 | 0.00385 | 0.49252 |
| 17 | AZT + ddc | 102 | 0 | 0.06875 | 0.00443 | 0.26651 | 0.00000 | 0.71206 | 0.00026 | 0.41889 |
| 18 | AZT + ddc | 180 | 1 | . | . | . | . | . | . | . |
| 19 | AZT + ddc + saquinivir | 0 | . | 1.00000 | 1.00000 | 1.00000 | . | . | . | . |
| 20 | AZT + ddc + saquinivir | 2 | 0 | 0.94118 | 0.65018 | 0.99150 | 0.00000 | 0.99981 | 0.26951 | 0.99720 |
| 21 | AZT + ddc + saquinivir | 3 | 0 | 0.88235 | 0.60598 | 0.96921 | 0.08457 | 0.99368 | 0.33241 | 0.98588 |
| 22 | AZT + ddc + saquinivir | 4 | 0 | 0.82353 | 0.54713 | 0.93941 | 0.21800 | 0.97556 | 0.31722 | 0.96770 |
| 23 | AZT + ddc + saquinivir | 12 | 0 | 0.76471 | 0.48828 | 0.90449 | 0.26286 | 0.94756 | 0.28564 | 0.94419 |
| 24 | AZT + ddc + saquinivir | 22 | 0 | 0.70588 | 0.43148 | 0.86560 | 0.26455 | 0.91280 | 0.24995 | 0.91622 |
| 25 | AZT + ddc + saquinivir | 48 | 0 | 0.64706 | 0.37715 | 0.82338 | 0.24619 | 0.87354 | 0.21398 | 0.88435 |
| 26 | AZT + ddc + saquinivir | 51 | 1 | 0.64706 | . | . | . | . | . | . |
| 27 | AZT + ddc + saquinivir | 56 | 1 | 0.64706 | . | . | . | . | . | . |
| 28 | AZT + ddc + saquinivir | 80 | 0 | 0.57516 | 0.30653 | 0.77204 | 0.20232 | 0.82576 | 0.16190 | 0.84534 |
| 29 | AZT + ddc + saquinivir | 85 | 0 | 0.50327 | 0.24358 | 0.71618 | 0.15461 | 0.77682 | 0.11914 | 0.80123 |
| 30 | AZT + ddc + saquinivir | 90 | 0 | 0.43137 | 0.18703 | 0.65595 | 0.10753 | 0.72832 | 0.08364 | 0.75209 |
| 31 | AZT + ddc + saquinivir | 94 | 1 | 0.43137 | . | . | . | . | . | . |
| 32 | AZT + ddc + saquinivir | 160 | 0 | 0.34510 | 0.12156 | 0.58442 | 0.05019 | 0.68501 | 0.04473 | 0.69468 |
| 33 | AZT + ddc + saquinivir | 171 | 0 | 0.25882 | 0.06909 | 0.50479 | 0.01346 | 0.65437 | 0.01952 | 0.62870 |
| 34 | AZT + ddc + saquinivir | 180 | 0 | 0.17255 | 0.02963 | 0.41588 | 0.00068 | 0.65473 | 0.00540 | 0.55360 |
| 35 | AZT + ddc + saquinivir | 180 | 1 | 0.17255 | . | . | . | . | . | . |
| 36 | AZT + ddc + saquinivir | 238 | 0 | 0.00000 | . | . | . | . | . | . |

*For arcsin-square root transformation, we have*

```
title2 "arcsin-square root transformation";
proc lifetest data=Q4
      outsurv = results3 conftype = asinsqrt confband = all
      plots = survival(CB = all);
    time time*status(0);
    by group;
run;

proc print data=results3;
run;
```




**confidence bands**
**arcsin-square root transformation**

| Obs | group | time | _CENSOR_ | SURVIVAL | SDF_LCL | SDF_UCL | HW_LCL | HW_UCL | EP_LCL | EP_UCL |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AZT + ddc | 0 | . | 1.00000 | 1.00000 | 1.00000 | . | . | . | . |
| 2 | AZT + ddc | 4 | 1 | 1.00000 | . | . | . | . | . | . |
| 3 | AZT + ddc | 6 | 0 | 0.93750 | 0.77210 | 0.99994 | 0.35193 | 1.00000 | 0.64045 | 1.00000 |
| 4 | AZT + ddc | 11 | 0 | 0.87500 | 0.67523 | 0.98652 | 0.42247 | 1.00000 | 0.53365 | 1.00000 |
| 5 | AZT + ddc | 12 | 0 | 0.81250 | 0.59204 | 0.95942 | 0.41114 | 0.99956 | 0.44731 | 0.99670 |
| 6 | AZT + ddc | 32 | 0 | 0.75000 | 0.51680 | 0.92437 | 0.37679 | 0.98126 | 0.37286 | 0.98235 |
| 7 | AZT + ddc | 35 | 0 | 0.68750 | 0.44730 | 0.88358 | 0.33320 | 0.94761 | 0.30699 | 0.95942 |
| 8 | AZT + ddc | 38 | 1 | 0.68750 | . | . | . | . | . | . |
| 9 | AZT + ddc | 39 | 0 | 0.61875 | 0.37311 | 0.83573 | 0.27724 | 0.90417 | 0.23819 | 0.92882 |
| 10 | AZT + ddc | 45 | 0 | 0.55000 | 0.30502 | 0.78270 | 0.21990 | 0.85691 | 0.17877 | 0.89105 |
| 11 | AZT + ddc | 49 | 0 | 0.48125 | 0.24230 | 0.72486 | 0.16318 | 0.80818 | 0.12752 | 0.84643 |
| 12 | AZT + ddc | 75 | 0 | 0.41250 | 0.18471 | 0.66225 | 0.10896 | 0.75972 | 0.08399 | 0.79495 |
| 13 | AZT + ddc | 80 | 0 | 0.34375 | 0.13234 | 0.59467 | 0.05982 | 0.71341 | 0.04828 | 0.73623 |
| 14 | AZT + ddc | 84 | 0 | 0.27500 | 0.08569 | 0.52154 | 0.02029 | 0.67209 | 0.02112 | 0.66938 |
| 15 | AZT + ddc | 85 | 0 | 0.20625 | 0.04587 | 0.44172 | 0.00019 | 0.64156 | 0.00412 | 0.59268 |
| 16 | AZT + ddc | 87 | 0 | 0.13750 | 0.01524 | 0.35280 | 0.00000 | 0.63736 | 0.00000 | 0.50254 |
| 17 | AZT + ddc | 102 | 0 | 0.06875 | 0.00008 | 0.24850 | 0.00000 | 0.72280 | 0.00000 | 0.38961 |
| 18 | AZT + ddc | 180 | 1 | . | . | . | . | . | . | . |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 19 | AZT + ddc + saquinivir | 0 | . | 1.00000 | 1.00000 | 1.00000 | . | . | . | . |
| 20 | AZT + ddc + saquinivir | 2 | 0 | 0.94118 | 0.78458 | 0.99995 | 0.34320 | 1.00000 | 0.66458 | 1.00000 |
| 21 | AZT + ddc + saquinivir | 3 | 0 | 0.88235 | 0.69251 | 0.98741 | 0.42444 | 1.00000 | 0.56241 | 1.00000 |
| 22 | AZT + ddc + saquinivir | 4 | 0 | 0.82353 | 0.61319 | 0.96213 | 0.42021 | 1.00000 | 0.47915 | 0.99631 |
| 23 | AZT + ddc + saquinivir | 12 | 0 | 0.76471 | 0.54123 | 0.92949 | 0.39158 | 0.98609 | 0.40679 | 0.98222 |
| 24 | AZT + ddc + saquinivir | 22 | 0 | 0.70588 | 0.47453 | 0.89159 | 0.35297 | 0.95582 | 0.34222 | 0.96034 |
| 25 | AZT + ddc + saquinivir | 48 | 0 | 0.64706 | 0.41205 | 0.84946 | 0.30958 | 0.91740 | 0.28392 | 0.93220 |
| 26 | AZT + ddc + saquinivir | 51 | 1 | 0.64706 | . | . | . | . | . | . |
| 27 | AZT + ddc + saquinivir | 56 | 1 | 0.64706 | . | . | . | . | . | . |
| 28 | AZT + ddc + saquinivir | 80 | 0 | 0.57516 | 0.33416 | 0.79855 | 0.24643 | 0.87020 | 0.21116 | 0.89710 |
| 29 | AZT + ddc + saquinivir | 85 | 0 | 0.50327 | 0.26385 | 0.74189 | 0.18376 | 0.82128 | 0.15012 | 0.85452 |
| 30 | AZT + ddc + saquinivir | 90 | 0 | 0.43137 | 0.20017 | 0.67973 | 0.12370 | 0.77259 | 0.09913 | 0.80450 |
| 31 | AZT + ddc + saquinivir | 94 | 1 | 0.43137 | . | . | . | . | . | . |
| 32 | AZT + ddc + saquinivir | 160 | 0 | 0.34510 | 0.12674 | 0.60563 | 0.05137 | 0.73247 | 0.04512 | 0.74528 |
| 33 | AZT + ddc + saquinivir | 171 | 0 | 0.25882 | 0.06684 | 0.52055 | 0.00501 | 0.70501 | 0.01137 | 0.67165 |
| 34 | AZT + ddc + saquinivir | 180 | 0 | 0.17255 | 0.02226 | 0.42192 | 0.00000 | 0.70731 | 0.00000 | 0.57955 |
| 35 | AZT + ddc + saquinivir | 180 | 1 | 0.17255 | . | . | . | . | . | . |
| 36 | AZT + ddc + saquinivir | 238 | 0 | 0.00000 | . | . | . | . | . | . |

Question #5:

**Solution 5.** *For part (a), we have the number of the subjects at risk as a function of age as following:*

| *age* | *entry* | *exit* | *risk set* ($Y_i$) | $d_i$ |
|---|---|---|---|---|
| 58 | 2 | 0 | 2 | 0 |
| 59 | 1 | 0 | 3 | 0 |
| 60 | 2 | 1 | 5 | 1 |
| 61 | 2 | 0 | 6 | 0 |
| 62 | 3 | 1 | 9 | 1 |
| 63 | 2 | 1 | 10 | 1 |
| 64 | 1 | 0 | 10 | 0 |
| 65 | 0 | 2 | 10 | 2 |
| 66 | 2 | 1 | 10 | 1 |
| 67 | 3 | 0 | 12 | 0 |
| 68 | 1 | 2 | 13 | 2 |
| 69 | 3 | 4 | 14 | 2 |
| 70 | 3 | 2 | 13 | 2 |
| 71 | 1 | 2 | 12 | 2 |
| 72 | 2 | 3 | 12 | 2 |
| 73 | 2 | 2 | 11 | 1 |
| 74 | 0 | 2 | 9 | 1 |
| 75 | 0 | 0 | 7 | 0 |
| 76 | 0 | 2 | 7 | 1 |
| 77 | 0 | 1 | 5 | 1 |
| 78 | 0 | 1 | 4 | 0 |
| 79 | 0 | 2 | 3 | 0 |
| 80 | 0 | 1 | 1 | 0 |
| 81 | 0 | 0 | 0 | 0 |

*For part (b):*

*The conditional survival function for a diabetic who has survived to age 60 is:*

$$S_{60}(t) = P(X > t | X \geq 60)$$

*So the estimate is:*

$$\hat{S}_{60}(t) = \prod_{60 \leq t_i \leq t} \Big[1 - \frac{d_i}{Y_i}\Big], t \geq 60$$

*we have:*

$$\begin{aligned}
\hat{S}_{60}(t) &= 1 - \frac{1}{5} = 0.8 \ \textit{if}\ 60 \leq t < 62 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(60) \cdot (1 - \frac{1}{9}) = 0.71 \ \textit{if}\ 62 \leq t < 63 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(62) \cdot (1 - \frac{1}{10}) = 0.64 \ \textit{if}\ 63 \leq t < 65 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(63) \cdot (1 - \frac{2}{10}) = 0.512 \ \textit{if}\ 65 \leq t < 66 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(65) \cdot (1 - \frac{1}{10}) = 0.4608 \ \textit{if}\ 66 \leq t < 68 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(66) \cdot (1 - \frac{2}{13}) = 0.39 \ \textit{if}\ 68 \leq t < 69 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(68) \cdot (1 - \frac{2}{14}) = 0.334 \ \textit{if}\ 69 \leq t < 70 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(69) \cdot (1 - \frac{2}{13}) = 0.28 \ \textit{if}\ 70 \leq t < 71 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(70) \cdot (1 - \frac{2}{12}) = 0.236 \ \textit{if}\ 71 \leq t < 72 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(71) \cdot (1 - \frac{2}{12}) = 0.196 \ \textit{if}\ 72 \leq t < 73 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(72) \cdot (1 - \frac{1}{11}) = 0.179 \ \textit{if}\ 73 \leq t < 74 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(73) \cdot (1 - \frac{1}{9}) = 0.159 \ \textit{if}\ 74 \leq t < 76 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(74) \cdot (1 - \frac{1}{7}) = 0.136 \ \textit{if}\ 76 \leq t < 77 \\
\hat{S}_{60}(t) &= \hat{S}_{60}(76) \cdot (1 - \frac{1}{5}) = 0.109 \ \textit{if}\ t \geq 77
\end{aligned}$$

*For part (c):*

*Similar to part (b), we have*

$$`S_{70}(t) = P(X > t | X \geq 60)$$

*So the estimate is:*

$$\hat{S}_{70}(t) = \prod_{70 \le t_i \le t} \left[1 - \frac{d_i}{Y_i}\right], t \ge 70$$

*We have:*

$$\hat{S}_{70}(t) = (1 - \frac{2}{13}) = 0.846 \textit{ if } 70 \le t < 71$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(70) \cdot (1 - \frac{2}{12}) = 0.7051 \textit{ if } 71 \le t < 72$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(71) \cdot (1 - \frac{2}{12}) = 0.5876 \textit{ if } 72 \le t < 73$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(72) \cdot (1 - \frac{1}{11}) = 0.534 \textit{ if } 73 \le t < 74$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(73) \cdot (1 - \frac{1}{9}) = 0.4748 \textit{ if } 74 \le t < 76$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(74) \cdot (1 - \frac{1}{7}) = 0.407 \textit{ if } 76 \le t < 77$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(76) \cdot (1 - \frac{1}{5}) = 0.3256 \textit{ if } t \ge 77$$

*For part (d), if we ignore the left truncation and simply treat the data as right censored, we would just simply ignore the entry time, and pretend that everyone starts from age 0. So the relationship between age and risk set becomes:*

| *age* | *exit* | *risk set* ($Y_i$) | *event* ($e_i$) |
|---|---|---|---|
| 60 | 1 | 30 | 1 |
| 62 | 1 | 29 | 1 |
| 63 | 1 | 28 | 1 |
| 65 | 2 | 27 | 2 |
| 66 | 1 | 25 | 1 |
| 68 | 2 | 24 | 2 |
| 69 | 4 | 22 | 2 |
| 70 | 2 | 18 | 2 |
| 71 | 2 | 16 | 2 |
| 72 | 3 | 14 | 2 |
| 73 | 2 | 11 | 1 |
| 74 | 2 | 9 | 1 |
| 76 | 2 | 7 | 1 |
| 77 | 1 | 5 | 1 |

*So we have the estimate for survival function conditional on age greater or equal to* 60*:*

$$\hat{S}_{60}(t) = (1 - \frac{1}{30}) = 0.97 \; \textit{if } 60 \leq t < 62$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(60) \cdot (1 - \frac{1}{29}) = 0.93 \; \textit{if } 62 \leq t < 63$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(62) \cdot (1 - \frac{1}{28}) = 0.9 \; \textit{if } 63 \leq t < 65$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(63) \cdot (1 - \frac{2}{27}) = 0.833 \; \textit{if } 65 \leq t < 66$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(65) \cdot (1 - \frac{1}{25}) = 0.8 \; \textit{if } 66 \leq t < 68$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(66) \cdot (1 - \frac{2}{24}) = 0.733 \; \textit{if } 68 \leq t < 69$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(68) \cdot (1 - \frac{2}{22}) = 0.667 \; \textit{if } 69 \leq t < 70$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(69) \cdot (1 - \frac{2}{18}) = 0.593 \; \textit{if } 70 \leq t < 71$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(70) \cdot (1 - \frac{2}{16}) = 0.5185 \; \textit{if } 71 \leq t < 72$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(71) \cdot (1 - \frac{2}{14}) = 0.444 \; \textit{if } 72 \leq t < 73$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(72) \cdot (1 - \frac{1}{11}) = 0.404 \; \textit{if } 73 \leq t < 74$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(73) \cdot (1 - \frac{1}{9}) = 0.359 \; \textit{if } 74 \leq t < 76$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(74) \cdot (1 - \frac{1}{7}) = 0.3078 \; \textit{if } 76 \leq t < 77$$
$$\hat{S}_{60}(t) = \hat{S}_{60}(76) \cdot (1 - \frac{1}{5}) = 0.246 \; \textit{if } t \geq 77$$

*Similarly, conditional on patient greater than or equal to age* 70*, we have*

$$\hat{S}_{70}(t) = 1 - \frac{2}{18} = 0.89 \; \textit{if } 70 \leq t < 71$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(70) \cdot (1 - \frac{2}{16}) = 0.78 \; \textit{if } 71 \leq t < 72$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(71) \cdot (1 - \frac{2}{14}) = 0.67 \; \textit{if } 72 \leq t < 73$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(72) \cdot (1 - \frac{1}{11}) = 0.606 \; \textit{if } 73 \leq t < 74$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(73) \cdot (1 - \frac{1}{9}) = 0.5387 \; \textit{if } 74 \leq t < 76$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(74) \cdot (1 - \frac{1}{7}) = 0.462 \; \textit{if } 76 \leq t < 77$$
$$\hat{S}_{70}(t) = \hat{S}_{70}(76) \cdot (1 - \frac{1}{5}) = 0.3694 \; \textit{if } t \geq 77$$

Question #6:

I confirm that I have read understood the proof of Greenwood's formula for variance.