# UNIVERSITY OF OSLO

## Faculty of Mathematics and Natural Sciences

| | |
|---|---|
| Examination in: | STK4290 — Parametric lifetime modeling |
| Day of examination: | Thursday 12 June 2014 |
| Examination hours: | 14.30 – 18.30 |

This problem set consists of 5 pages.

| | |
|---|---|
| Appendices: | None |
| Permitted aids: | Approved calculator |

Please make sure that your copy of the problem set is complete before you attempt to answer anything.

## Problem 1 The Log-Laplace distribution

Let $\Psi(w)$ and $\psi(w)$ be, respectively, the cumulative distribution function and the probability density function for a random variable $W$ with range $(-\infty, \infty)$. Suppose that $\psi(w) > 0$ for all $w \in (-\infty, \infty)$.

### 1a

Explain how $\Psi(w)$ and $\psi(w)$ are used to define a *log-location-scale* family of distributions for a lifetime $T$, indexed by parameters $-\infty < \mu < \infty$ and $\sigma > 0$. What are the two parameters called?

Find the cumulative distribution function $F(t)$ and the probability density $f(t)$ for $T$ expressed by the functions $\Psi$ and $\psi$, and the parameters $\mu$ and $\sigma$.

In the rest of the problem it is assumed that the distribution of $W$ is given by a standard *Laplace distribution* (also called the *double exponential distribution*), i.e. it is assumed that

$$\psi(w) = \frac{1}{2}e^{-|w|} \text{ for } -\infty < w < \infty. \tag{1}$$

In the following you may use, without proving it, that the cumulative distribution function of $W$ is

$$\Psi(w) = \begin{cases} \frac{1}{2}e^{w} & \text{for } w \leq 0 \\ 1 - \frac{1}{2}e^{-w} & \text{for } w > 0 \end{cases}$$

and that the moment generating function for $W$ is

$$M_W(s) \equiv E(e^{sW}) = \frac{1}{1 - s^2} \text{ for } |s| < 1.$$

(Continued on page 2.)

### 1b

The resulting family of distributions for $T$ is called the *Log-Laplace distribution.* Show that this distribution can be reparameterized by $\alpha > 0, \theta > 0$ so that the cumulative distribution function can be written

$$F(t) = \begin{cases} \frac{1}{2}\left(\frac{t}{\theta}\right)^{\alpha} & \text{for } 0 \leq t \leq \theta, \\ 1 - \frac{1}{2}\left(\frac{t}{\theta}\right)^{-\alpha} & \text{for } t > \theta. \end{cases} \tag{2}$$

Express $\alpha$ and $\theta$ as functions of $\mu$ and $\sigma$.

### 1c

Find expressions for the density $f(t)$ and the hazard rate $z(t)$ for $T$ under the parameterization (2).

Sketch the functions $f(t)$ and $z(t)$ when $\alpha = \theta = 1$.

### 1d

Find expressions for the expectation and variance of the distribution (2). Show that the expectation is defined only if $\alpha > 1$, while the variance is defined only if $\alpha > 2$.

(*Hint:* You may, for example, first use the moment generating function of $W$ to find a formula for $E(T^r)$ for general $r$.)

### 1e

Show that the median of the distribution (2) is $\theta$.

Let $t_p$ be the $100p$-percentile of the distribution (2), such that $F(t_p) = p$, for $0 < p < 1$. Find $t_p$ as a function of $p$.

(*Hint:* Distinguish between the cases $p \leq 1/2$ and $p > 1/2$.

## Problem 2 TTT-plot

Let $T_1, \ldots, T_n$ be independent (and uncensored) observations of a lifetime $T$ with cumulative distribution function $F(t)$ and density $f(t)$. It is assumed that $f(t) > 0$ for all $t > 0$, so that $F(t)$ is everywhere strictly increasing. It is also assumed that $\mu = E(T)$ exists.

Let the ordered lifetimes be

$$T_{(1)} < T_{(2)} < \cdots < T_{(n)}.$$

The empirical distribution function $F_n(t)$ is defined by

$$F_n(t) = \frac{i}{n} \text{ for } T_{(i)} \leq t < T_{(i+1)};\ i = 0, 1, \ldots, n,$$

where $T_{(0)} = 0,\ T_{(n+1)} = \infty$.

*(Continued on page 3.)*

## 2a

Let $\mathcal{T}(t)$ be the "Total Time on Test" until (and including) time $t$, for $t > 0$.

Explain briefly why we can write, for $i = 1, \ldots, n$,

$$\mathcal{T}(T_{(i)}) = nT_{(1)} + (n-1)(T_{(2)} - T_{(1)}) + \ldots + (n-i+1)(T_{(i)} - T_{(i-1)}).$$

Can you find a simpler expression for $\mathcal{T}(T_{(n)})$?

Show that for each $i = 1, 2, \ldots, n$ we have

$$\frac{1}{n}\mathcal{T}(T_{(i)}) = \int_0^{F_n^{-1}(i/n)} (1 - F_n(t))dt.$$

(*Hint:* For a general cumulative distribution function $H(x)$ is defined

$$H^{-1}(u) = \inf\{x : H(x) \geq u\} \text{ for } 0 \leq u \leq 1,$$

so that $F_n^{-1}(i/n) = T_{(i)}$ for $i = 1, \ldots, n$.)

## 2b

Assume that $n \to \infty$ and at the same time $i \to \infty$ in such a way that $i/n \to u$, where $0 \leq u \leq 1$. It can be shown (you should not do it) that we have the convergence

$$\int_0^{F_n^{-1}(i/n)} (1 - F_n(t))dt \to \int_0^{F^{-1}(u)} (1 - F(t))dt \equiv G_F(u)$$

with probability 1, uniformly in $0 \leq u \leq 1$.

The TTT-plot based on the observations $T_1, \ldots, T_n$ is known to be given by the points

$$\left(\frac{i}{n}, \frac{\mathcal{T}(T_{(i)})}{\mathcal{T}(T_{(n)})}\right), \; i = 1, \ldots, n.$$

Explain that the TTT-plot, when $n \to \infty$ and $i/n \to u$, converges to the curve defined by

$$\left(u, \frac{G_F(u)}{\mu}\right), \quad 0 \leq u \leq 1.$$

## 2c

Assume in this subpoint that $T$ is exponentially distributed with mean $\theta > 0$, i.e. $T$ has cumulative distribution function

$$F(t) = 1 - e^{-t/\theta} \text{ for } t > 0.$$

Calculate the function $G_F(u)$. To which curve will the TTT-plot thus converge when the observations come from an exponential distribution?

Why is this a useful result for practical use of the TTT-plot?

*(Continued on page 4.)*

### 2d

Let again $T$ have a general distribution. Show by differentiating $G_F(u)$ that

$$G_F'(F(t)) = \frac{1}{z(t)} \text{ for } t > 0, \tag{3}$$

where $z(t)$ is the hazard rate of $T$.

(*Hint:* You may use the formula for differentiation of the inverse of a function $k(x)$:

$$[k^{-1}]'(x) = \frac{1}{k'(k^{-1}(x))} \text{ ).}$$

Explain how the result (3) can be used to interpret the TTT-plot with respect to increasing or decreasing hazard rate.

## Problem 3 Repairable system

We consider a repairable system that operates from time $t = 0$. At failures, the system is repaired and put back into operation immediately.

Let $N(t)$ for $t > 0$ denote the number of failures in the time interval $(0, t]$. It is assumed that $N(t)$ is a non-homogeneous Poisson process (NHPP) with intensity function $w(t)$ and cumulative intensity function $W(t) = \int_0^t w(u)du$.

### 3a

What are the characteristics defining $N(t)$ as an NHPP?

Explain briefly what it means that a repair is, respectively, minimal and perfect. Which of these two types of repairs are modeled by NHPP? Give an explanation.

### 3b

Suppose now that the functions $w(t)$ and $W(t)$ are modeled by parametric functions $w(t;\theta)$ and $W(t;\theta)$, and that one would like to make statistical inference about $\theta$.

The system is observed in the time interval $[0, \tau]$, where one only registers the number of failures occuring in disjoint time intervals defined by the grid

$$h_0 = 0 < h_1 < h_2 < \cdots < h_r = \tau.$$

Let $D_i$ be the number of observed failures in the interval $(h_{i-1}, h_i]$, for $i = 1, 2, \ldots, r$. Explain why $D_1, \ldots, D_r$ are independent and Poisson-distributed, with

$$E(D_i) = W(h_i;\theta) - W(h_{i-1};\theta).$$

Use this to show that the likelihood function for $\theta$ when the observations are given as $d_1, \ldots, d_r$, can be written

$$L(\theta; d_1, \ldots, d_r) = \left\{ \prod_{i=1}^{r} \frac{[W(h_i;\theta) - W(h_{i-1};\theta)]^{d_i}}{d_i!} \right\} e^{-W(\tau;\theta)}. \tag{4}$$

*(Continued on page 5.)*

**3c**

Suppose finally that one still has registered the exact failure times $s_1, s_2, \ldots, s_n$ in the interval $[0, \tau]$.

Show how the likelihood function in (4) leads to the following expression for the likelihood function for this case:

$$L(\theta; s_1, \ldots, s_n) = \left\{ \prod_{\ell=1}^{n} w(s_\ell; \theta) \right\} e^{-W(\tau;\theta)}.$$

END