Name Jerry

Student Number

STA 312 f2023 Quiz 4

1. (3 points) The expected value of a Weibull random variable T with parameters α and λ is $E(T) = \frac{1}{\lambda}\Gamma\left(\frac{1}{\alpha}+1\right)$. You don't have to show this. To obtain the standard error of *estimated* $E(T)$ for your R work, you needed to calculate $\dot{g}(\alpha,\lambda)$. Show the calculation of $\dot{g}(\alpha,\lambda)$ in the space below. **Circle your final answer.**

$$\dot{g}(\alpha,\lambda) = \left(\frac{\partial g}{\partial \alpha}, \frac{\partial g}{\partial \lambda}\right)$$

$$= \left(\frac{\partial}{\partial\alpha}\frac{1}{\lambda}\Gamma(\alpha^{-1}+1), \frac{\partial}{\partial\lambda}\lambda^{-1}\Gamma\left(\frac{1}{\alpha}+1\right)\right)$$

$$= \left(\frac{1}{\lambda}\Gamma'\left(\frac{1}{\alpha}+1\right)(-1)\alpha^{-2}, (-1)\lambda^{-2}\Gamma\left(\frac{1}{\alpha}+1\right)\right)$$

$$= \left(-\frac{\Gamma'\left(\frac{1}{\alpha}+1\right)}{\alpha^2\lambda}, -\frac{\Gamma\left(\frac{1}{\alpha}+1\right)}{\lambda^2}\right)$$

2. (2 points) For Question 1 of Assignment 4, you analyzed numerical data from a Weibull distribution, and you produced a 95% confidence interval for $E(T)$. Write the confidence interval in the space below: Just two numbers. On your printout, circle the numbers and write "Question 2" beside them. **The code that produced the confidence interval for $E(T)$ must be shown.**

(3.44, 3.79)

3. (5 points) Let T be a continuous random variable with $P(T > 0) = 1$, density $f(t)$ and cumulatve distribution function $F(t) = P(T \leq t)$. Prove $h(t) = \frac{f(t)}{S(t)}$. You may use anything on the formula sheet except the fact you are proving.

$$h(t) \stackrel{def}{=} \lim_{\Delta \to 0} \frac{P(t < T < t+\Delta \mid T > t)}{\Delta}$$

$$= \lim_{\Delta \to 0} \frac{P(t < T < t+\Delta, T > t)}{\Delta \, P(T > t)}$$

$$= \lim_{\Delta \to 0} \frac{P(t < T < t+\Delta)}{\Delta \, S(t)}$$

$$= \frac{1}{S(t)} \lim_{\Delta \to 0} \frac{F(t+\Delta) - F(t)}{\Delta}$$

$$= \frac{1}{S(t)} f(t) = \frac{f(t)}{S(t)} \quad \square$$

Please attach the printout with your answer to Question 2 of this quiz (Question 1d of the assignment). **The code that produced the confidence interval for $E(T)$ must be shown.** Make sure your name and student number are written on the printout.

STA312f23 Assignment 3 Q3e Printout

```
> # Assignment 3, Question 3e.
> rm(list=ls()); options(scipen=999)
> x = scan("http://www.utstat.toronto.edu/brunner/data/legal/Weibull.data1.txt")
Read 500 items
>
> # (i) Find MLE
> mloglike = function(theta,datta)
+     {
+     alpha = theta[1]; lambda = theta[2]
+     n = length(datta)
+     value = lambda^alpha*sum(datta^alpha) -
+             n*log(alpha) - n*alpha*log(lambda) - (alpha-1)*sum(log(datta))
+     return(value)
+     } # End of function mloglike
>
> # Testing
> mloglike(c(2,0.25),datta=x) # 1019.647
[1] 1019.647
> -sum(dweibull(x,shape=2,scale = 4,log=TRUE)) # 1019.647
[1] 1019.647
>
> # Find MLE: Truth is alpha = 2 and lambda=1/4
> startvals = c(1,1)
> search1 = optim(par=startvals, fn=mloglike, datta=x,
+                 hessian=TRUE, lower=c(0,0), method='L-BFGS-B')
> search1
$par
[1] 1.9124466 0.2454345

$value
[1] 1018.006

$counts
function gradient
      13       13

$convergence
[1] 0

$message
[1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"

$hessian
         [,1]       [,2]
[1,] 244.4778   843.0713
[2,] 843.0713 30358.9099

>
> alphahat = search1$par[1]; alphahat # Truth is 2
[1] 1.912447
> lambdahat = search1$par[2]; lambdahat # Truth is 1/4
[1] 0.2454345
> Vhat = solve(search1$hessian); Vhat
              [,1]           [,2]
[1,]  0.0045235447 -0.00012561949
[2,] -0.0001256195  0.00003642773
>
```

```
>
> # (ii) 95% CI for alpha
>
> # CI for alpha
> se_alphahat = sqrt(Vhat[1,1])
> lower95 = alphahat - 1.96*se_alphahat; upper95 = alphahat + 1.96*se_alphahat
> c(lower95,upper95)
[1] 1.780622 2.044271
>
> # (iii) Point estimate of expected value
>
> # Give a point estimate of the expected value (mu = gamma(1+1/alpha)/lambda).
Truth is 3.544908
> muhat = gamma(1+1/alphahat)/lambdahat; muhat
[1] 3.614746
> mean(x) # Compare
[1] 3.618727
>
> # (iv) 95% CI for mu
> # Hint is help(digamma) for gdot
>
> gprime = digamma(1+1/alphahat)*gamma(1+1/alphahat)
> gdot = cbind( -gprime/(alphahat^2*lambdahat), -gamma(1+1/alphahat)/lambdahat^2)
> v_muhat = as.numeric( gdot %*% Vhat %*% t(gdot) ); se_muhat = sqrt(v_muhat)
> lower95 = muhat - 1.96*se_muhat; upper95 = muhat + 1.96*se_muhat
> c(lower95,muhat,upper95)
[1] 3.442696 3.614746 3.786796
```

Question 2

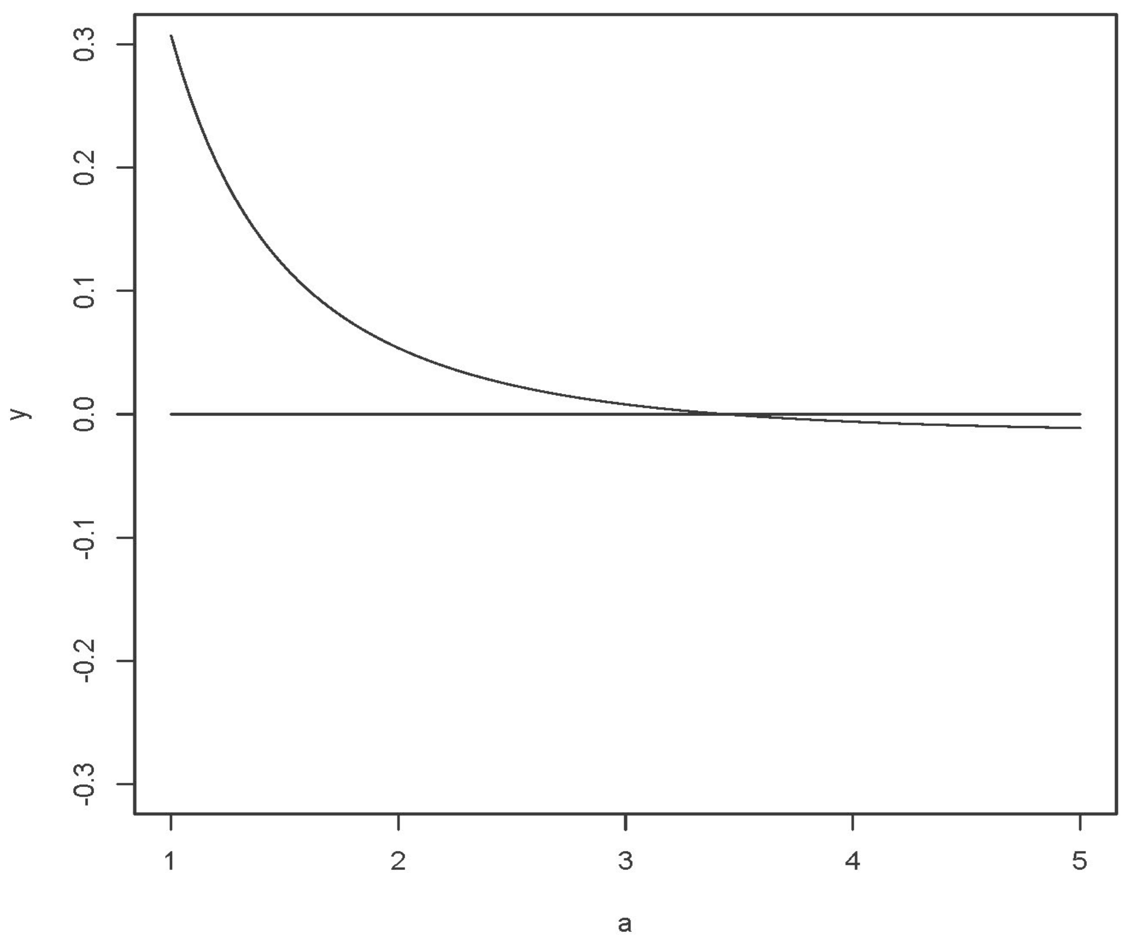
```
>
> t.test(x) # For comparison

      One Sample t-test

data:  x
t = 41.284, df = 499, p-value < 0.00000000000000022
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 3.446509 3.790946
sample estimates:
mean of x
 3.618727


>
> # (v) Estimate median
>
> mhat = 1/lambdahat * log(2)^(1/alphahat); mhat # Truth is 3.544908
[1] 3.363827
> median(x) # Sample median
[1] 3.411699
>
> # (vi) 95% CI for median
> # D[b^(1/a),a] works in Wolfram Alpha, hand as a check.
>
> gdot2 = cbind( - log(2)^(1/alphahat)*log(log(2))/(lambdahat*alphahat^2),
+                - log(2)^(1/alphahat)/lambdahat^2 )
> v_mhat = as.numeric( gdot2 %*% Vhat %*% t(gdot2) ); se_mhat = sqrt(v_mhat)
> lower95 = mhat - 1.96*se_mhat; upper95 = mhat + 1.96*se_mhat
> c(lower95,upper95)
[1] 3.182938 3.544715
>
> # > ci.median(x) # From asbio package
> # 95% Confidence interval for population median
> # Estimate     2.5%    97.5%
> # 3.411699 3.176731 3.624114
>
```

```
>
> # (vii) Test mean = median
>
> # Straightforward delta method.
> gdot3 = gdot-gdot2 # Derivaltive of a difference is difference of derivatives.
> v_diff1 = as.numeric( gdot3 %*% Vhat %*% t(gdot3) ); se_diff1 = sqrt(v_diff1)
>
> Z1 = (muhat-mhat)/se_diff1;  Z1
[1] 9.923228
>
> # The other methods simplify the null hypothesis, and test a simple
> # equivalent statement.
> # I believe gamma(1+1/alpha) = log2^(1/alpha) iff alpha has a particular
> # numerical value,
> # the root of g(alpha) = gamma(1+1/alpha) - log2^(1/alpha)
>
> # Plot it to see if it has a root.
> a = seq(from=1,to=5,length=101)
> y = gamma(1+1/a) - log(2)^(1/a)
> plot(a,y,type='l', ylim=c(-0.3,.3))
> xx = c(1,5); yy = c(0,0); lines(xx,yy)
> # Root is between alpha=3 and alpha=4
```



```
>
> g = function(x) gamma(1+1/x) - log(2)^(1/x)
> intersection = uniroot(g,c(3,4)); intersection
$root
[1] 3.439545

$f.root
[1] -0.00000006785751

$iter
[1] 5

$init.it
[1] NA

$estim.prec
[1] 0.00006103516

> # Already see reject H0 because it's outside 95% CI for alpha
> alpha0 = intersection$root; alpha0
[1] 3.439545
>
> Z2 = (alphahat-alpha0)/se_alphahat; Z2
[1] -22.70532
>
> # LR test of H0: alpha=alpha0. Need restricted MLE lambdahat0
>
> restmll = function(lambda,datta) # Restricted minus loglike
+     {
+     alpha = alpha0
+     n = length(datta)
+     value = lambda^alpha*sum(datta^alpha) -
+             n*log(alpha) - n*alpha*log(lambda) - (alpha-1)*sum(log(datta))
+     if(value>10^10) value = 10^10
+     return(value)
+     } # End of function restmll
>
> # Try some values
>
> restmll(.25,datta=x)
[1] 1305.871
> restmll(1,datta=x)
[1] 101511.3
> restmll(.001,datta=x)
[1] 9922.515
> # That brackets it
>
> search2 = optim(par=lambdahat, fn=restmll, datta=x,
+                 lower=0, upper=1, method='L-BFGS-B')
> search2
$par
[1] 0.2121786

$value
[1] 1208.953

$counts
function gradient
       6        6

$convergence
[1] 0

$message
[1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"
```

```
>
> # G-squared is twice difference between minus log likelihoods
> # And of course the minus LL (lack of fit) is greater for the restricted model.
>
> Gsq = 2 * (search2$value - search1$value); Gsq
[1] 381.8942
>
> round(c(Z1,Z2,Gsq),2)
[1]   9.92 -22.71 381.89
> Z2^2 # The Wald statistc for H0: alpha=alpha0
[1] 515.5317
```