Antonio Osamu Katagiri Tanaka A01212611

1. log-rank test would be implemented when comparing survival times of two treatments with time series censored data.

the null hypothesis Ho to test should be: $H_0(t): h_{A,t} = h_{B,t}$

↳ meaning that the hazard rate for time t is the same for both treatments.

* the purpose is to answer: If n_d subjects are randomly drawn from n, what is the probability distribution for the number of drawn subjects to be from treatment A?

↳ where n_d = total no. of deaths/failures

n_s = total no. of survivals/passes

n_A = total no. of subjects in treatment A

n_B = total no. of subjects in treatment B

$n = n_d + n_s = n_A + n_B$

$\boxed{\text{step 1}}$ compute the log-rank statistic

$$z = \frac{\sum_{i=1}^{n}(y_i - E_i)}{\left(\sum_{i=1}^{n} V_i\right)^{1/2}} \qquad \text{with: } E(y) = \frac{n_A n_d}{n}, \quad V(y) = \frac{n_A n_B n_d n_s}{n^2(n-1)}$$

$\boxed{\text{step 2}}$ compare z with normal critical values to get the statistic conclusion.

2. The likelihood of $\underline{y}$ is: $f_{\underline{\lambda}}(y) = e^{\underline{\lambda}\cdot\underline{y} - \gamma(\underline{\lambda})} \cdot f_0(y)$

↳ For a binomial distribution $\lambda = \log\left(\frac{\pi}{1+\pi}\right)$ & $\gamma(\lambda) = n\log(1+e^{\lambda})$; substitution gives:

$f_\lambda(y) = \left[\left(\frac{\pi}{1+\pi}\right)y - (1+e^{\lambda})n\right] f_0(y)$

$\lambda = \frac{\frac{\pi}{1+\pi}y}{n\left(\frac{\pi}{1+\pi}\right)}$

the generalization of the deviance function between f_1 & f_2 is:

$$D(f_1,f_2) = 2\cdot\int_{\text{sample space } y} f_1(y)\cdot\log\frac{f_1(y)}{f_2(y)}\,dy = 2\int_y \frac{\pi_1}{1+\pi_1}y \div n\left(\frac{\pi_1}{1+\pi_1}\right)\log\left(\frac{\frac{\pi_1}{1+\pi_1}y \div n\left(\frac{\pi_1}{1+\pi_1}\right)}{\frac{\pi_2}{1+\pi_2}y \div n\left(\frac{\pi_2}{1+\pi_2}\right)}\right)dy$$

$$D(f_1,f_2)_{Bin} = 2n\left(\pi_1\log\left(\frac{\pi_1}{\pi_2}\right) + (1-\pi_1)\log\left(\frac{1-\pi_1}{1-\pi_2}\right)\right)$$

$\boxed{\text{3a}}$ TRUE.

- James-Stein gets observations from Normal distributions which have different means for each observation
- extreme shrinkage is to set each observation as the average of all observations, void shrinkage is to set each observation as its own average. James-Stein's in between.

$\boxed{3b}$ FALSE

$\cdot \int e^{-\frac{1}{2}(z^2)} dz = \int e^{-z} dz = -e^{-z} \Rightarrow \left[-e^{-z}\right]_{-\infty}^{\infty} \neq \sqrt{2\pi}$

$\boxed{3c}$ FALSE

MLE can cause overfitting identification problems in high dimensions; if a lot of parameters in θ are fitted, the fit would become very specific to the current data set and would not represent the population.

$\boxed{3d}$ FALSE

In Ridge Regression, if the β coefficients are small, then we get a better response. Making the coefficients small make the variance to decrease, but it introduces some bias, therefore the β coefficients are to be large enough but no more than necesary.

$\boxed{3e}$ TRUE

In Bootstrap each observation on each sample is to be randomly drawn with equal probability and with replacement from the initial sample.
The resampling is to be the same as Bootstrap replaces an unknown distribution with an estimate of it.

$\boxed{4}$ a life time is represented by X; so $f_i = Pr(X=i)$ is the probability of dying at age i and $S_i = \sum_{j \geq i} f_j = Pr\{X \geq i\}$ is the probability of surviving past age $i-1$. The hazard rate at age i is $h_i = f_i / S_i = Pr\{X=i \mid X \geq i\}$. The probability of surviving past age j given survival past age $i-1$ is $S_{ij} = \prod_{k=i} (1-h_k) = Pr(X>j \mid X \geq i)$ ∮ $\hat{S}_j = \prod_{k \leq i} \left(\frac{n-k}{n-k+1}\right)^{dk}$ is the Kaplan Meier Estimate.

$\boxed{5}$ by "bootstrap resampling plan" reasoning, $K_b = [K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_8, K_9, K_{10}]$

$K_b = [3\ 0\ 0\ 0\ 1\ 1\ 2\ 1\ 1\ 1]$

as K_p follows a multinomial distribution,

$f(K_b) = \frac{10!}{3!\,(1!)^5\,2!} \cdot \frac{1}{10^{10}} = 3.024 \times 10^{-3}$ % probability of getting that particular sample.