Quiz → Level 1

1) → $X_i \overset{iid}{\sim} Exp(\lambda = 1/24)$

hazard rates are constant ⇒ $|h_1(x_1) - h_2(x_2)| = 0$

---

2) → $X \sim Exp(\lambda)$

$V(X) = 1/9 \Rightarrow \lambda = 3$

→ median of X → $0.5 = F_X(m)$

$= 1 - e^{-3m}$

$\Rightarrow m = 0.231$

---

3) → $X \sim Normal(\mu, \sigma^2 = 64)$

$H_0: \mu = 2$

$H_A: \mu = 6$

RR: $\bar{x} > 5.25$

→ Type II error rate = 0.10 → $0.10 = P(\bar{X} < 5.25 \mid \mu = 6)$

$= P(Z < 1.281)$

$1.281 = \frac{5.25 - 6}{8/\sqrt{n}}$

$n = \left[\frac{1.281(8)}{(5.25-6)}\right]^2$

$= 186.70 \to 187$

---

4) $\ln(\lambda) = x^T\beta$

$\Rightarrow \lambda = e^{x^T\beta} \Rightarrow V(Y) = \lambda = e$ plug in x & $\hat{\beta}$

---

5) Qualitative

Quiz → Level 3

1) Lasso → as λ increases, more shrinkage occurs

---

2) Logit → $p = \frac{e^{x^T\beta}}{1+e^{x^T\beta}}$ $= 0.772$ } Probit → $p = \Phi^{-1}(x^T\beta)$ $= 0.884$ } $|logit - probit| = 0.112$

---

3) Poisson w/ log link → $\lambda = e^{x^T\beta}$ ⇒ multiply by pop size ×100,000 $= 6.44$

---

4) Qualitative

---

5) $x_i \le x$: $\hat{y}_i$ = $x_i > x$: $\hat{y}_i$

$17.725 + 10x + a x^2 = 310.83 - 6x + 0.73x^2$ c @ $x = 4.7$

$64.725 + 22.09a = 294.4252 - 4.26$

① $22.09a + 4.26 = 230.2002$ c @ $x = 33$

② $1089a + 33b = 747.185$

$$\left[\begin{array}{cc|c} 1 & \frac{33}{1089} & \frac{747.185}{1089} \\ 22.09 & 4.7 & 230.2002 \end{array}\right] r_2 - 22.09r_1$$

$$\left[\begin{array}{cc|c} 1 & \frac{33}{1089} & \frac{747.185}{1089} \\ 0 & 4.0306 & 215.36 \end{array}\right]$$

$\Rightarrow b = 53.33$

$\Rightarrow a = -0.93$

problems aren't that hard!

→ find a → $17.725 + 10(4.7) + a(4.7^2) = 53.68$

$\Rightarrow a = -0.5$

→ find b → c ... ? $\Rightarrow b = 22.82$

→ First derivatives must be equivalent at x

$\Rightarrow 2ax + 10 = 2(0.73)x - b$

($a \to -0.5$, $b \to 22.82$)

$\Rightarrow x = 15.5$

Quiz → Level 4

→ difference in SSE (additional reduction in SSE), relative to full model SSE

Q1) $F = \frac{(SSE_r - SSE_F)/(df_r - df_F)}{SSE_F/df_F} = \frac{(2731 - 2104)/2}{2104/25} = 4.022$

p-value $= P(F > 16.322) = 0.03$, $F_{2,25}$

★ double check letters formula + its formula!!!

Q2) → residual st error $= \sqrt{MSE}$

→ $y_4 = 27$, $\hat{y}_4 = 23.6394$ ⇒ $e_i = 3.3606$

→ $e_{4,std} = \frac{e_4}{\sqrt{1-h_4}} = \frac{e_i}{\sqrt{MSE(1-h_i)}} = 0.7936$ } → $\frac{3.3606}{\sqrt{MSE(1-h_i)}} = 0.7936$

$MSE(1-h_i) = 17.932$

→ $D_i = \frac{e_{std,4}^2 h_i}{(p+1)(1-h_i)} = 0.15403$ } → $\frac{0.7936^2 h_i}{(4)(1-h_i)} = 0.15403$

$\frac{h_i}{1-h_i} = 0.9782$

$1.9782 h_i = 0.9782$

$h_i = 0.4945$ ⇒ $MSE = 35.473$

⇒ $\sqrt{MSE} = 5.956$

Q3) Ask # of models → forward stepwise →($p=10$) $10 + 9 + 8 + \dots + 1 + 1 = 56$

→ best subset → 2^{10}

$2^n = \sum_{i=0}^{n} \binom{n}{i}$

Q4) $D = 2[\ell(\hat{\beta}_{full}) - \ell(\beta)]$

$= 2[\ln(0.00001851) - \ln(0.000001244)]$

$= 5.4$

$D \sim \chi^2_{\text{difference in \# params estimated}}$

Q5) → $p = 6$, $n = 100$

$\ell(\hat{\pi}) = -346.7$ → fitted model

$\ell(\tilde{\pi}) = -361.5$ → null model

→ $D = 2[\ell(\hat{\pi}) - \ell(\tilde{\pi})] = 29.6$ ✓

$= LRT \sim \chi^2_5$

→ $AIC = -2\ell(\hat{\pi}) + 2p = 705.4$

$BIC = -2\ell(\hat{\pi}) + \ln(n) \cdot p = 721.03$

→ Pseudo $R^2 = \frac{\ell_{null} - \ell(\hat{\beta})}{\ell_{null}} = 1 - \frac{\ell(\hat{\beta})}{\ell_{null}} = 0.0409$