Q4 Solutions

(a) The joint posterior of (P_1, P_2) is

$$\pi(P_1, P_2 \mid X, y) \propto P(X, y \mid P_1, P_2)\, \pi(P_1, P_2)$$

$$\propto P_1^{X} P_2^{y} (1-P_1)^{n_1 - X} (1-P_2)^{n_2 - y}$$

$$= P_1^{X+1-1} (1-P_1)^{n_1 - X+1-1} P_2^{y+1-1} (1-P_2)^{n_2 - y+1-1}$$

Thus, a posteriori, (P_1, P_2) are independently distributed as beta distributions with

$$P_1 \mid X \sim Beta(X+1, n_1 - X + 1)$$
$$P_2 \mid y \sim Beta(y+1, n_2 - y + 1)$$

Under squared error loss, the Bayes rule (estimator) of $P_1 - P_2$ is the posterior mean, so that the Bayes rule is

$$E[(P_1 - P_2) \mid X, y]$$

$$= E[P_1 \mid X] - E(P_2 \mid y)$$

$$= \frac{X+1}{X+1+n_1-X+1} - \frac{y+1}{y+1+n_2-y+1} = \boxed{\frac{X+1}{n_1+2} - \frac{(y+1)}{n_2+2}}$$

(b) The frequentist risk is given by

$R(\theta, d_\Lambda) = E_{X,y|\theta}(\theta - d_\Lambda)^2$, where $\theta = P_1 - P_2$

and $d_\Lambda = \frac{X+1}{n_1+2} - \frac{(y+1)}{n_2+2}$

$$E(\theta - d_\Lambda)^2 = (P_1 - P_2)^2 - 2(P_1 - P_2)E(d_\Lambda) + E(d_\Lambda^2)$$

$$= (P_1 - P_2)^2 - 2(P_1 - P_2)E(d_\Lambda) + Var(d_\Lambda) + (E(d_\Lambda))^2$$

$$= (P_1 - P_2)^2 + E(d_\Lambda)\left[E(d_\Lambda) - 2(P_1 - P_2)\right] + Var(d_\Lambda).$$

$$E(d_\Lambda) = \frac{E(X)+1}{n_1+2} - \frac{E(Y)+1}{n_2+2}$$

$$= \frac{n_1 P_1 + 1}{n_1+2} - \frac{n_2 P_2 + 1}{n_2+2}$$

$$Var(d_\Lambda) = \frac{1}{(n_1+2)^2} Var(X) + \frac{1}{(n_2+2)^2} Var(Y)$$

$$= \frac{n_1 P_1 (1 - P_1)}{(n_1+2)^2} + \frac{n_2 P_2 (1 - P_2)}{(n_2+2)^2}$$

(b) Cont'd

The Bayes risk is

$$\mathcal{R} = \int R(\theta, d_\Lambda)\, \pi(\theta)\, d\theta$$

$$\int_0^1\int_0^1 \left[(P_1-P_2)^2 - 2(P_1-P_2)E(d_\Lambda) + Var(d_\Lambda) + (E(d_\Lambda))^2\right] dP_1 dP_2$$

$$= Var(d_\Lambda) + (E(d_\Lambda))^2 + \int_0^1\int_0^1 (P_1-P_2)^2\, dP_1\, dP_2$$

$$= Var(d_\Lambda) + (E(d_\Lambda))^2 + \frac{1}{6}$$

Note that $-2E(d_\Lambda)\int_0^1\int_0^1 (P_1-P_2)\, dP_1 dP_2 = 0$

(b) Cont'd

Is the Bayes rule admissible?

The Bayes rule just derived has a finite Bayes risk, and Bayes rules with finite Bayes risk are unique. Thus, any unique Bayes rule with finite Bayes risk is admissible. Thus d_Λ is admissible.

(c) We can write the joint distribution of X and Y as

$$P(X=x, Y=y) = \binom{n_1}{x}\binom{n_2}{y} p_1^{x} p_2^{y} (1-p_1)^{n_1-x}(1-p_2)^{n_2-y}$$

$$= \binom{n_1}{x}\binom{n_2}{y}(1-p_1)^{n_1}(1-p_2)^{n_2}$$

$$\cdot \exp\left[y\left(\log\left(\frac{p_2}{1-p_2}\right) - \log\left(\frac{p_1}{1-p_1}\right)\right) + (x+y)\log\left(\frac{p_1}{1-p_1}\right)\right].$$

Then, we can apply Theorem 2.7 of the notes to obtain the UMPU test for the hypotheses $H_0: p_1 \le p_2$ vs. $H_1: p_1 > p_2$

using $\theta = \log\left(\dfrac{\frac{p_2}{1-p_2}}{\frac{p_1}{1-p_1}}\right) = \log(\rho)$

$$\rho = \log\left(\frac{\frac{p_2}{1-p_2}}{\frac{p_1}{1-p_1}}\right)$$

$U = Y$, $T = X+Y$.

Thus, the hypothesis is equivalent to

$H_0: \theta \geq 0$ vs. $H_1: \theta < 0$

and the rejection region is given by

$$\phi(U) = \begin{cases} 1 & \text{if } U < c(t) \\ \gamma(t) & \text{if } U = c(t) \\ 0 & \text{if } U > c(t) \end{cases}$$

$c(t)$ is found by solving $\alpha = E_{\theta=0}[\phi(U) \mid T=t]$

To find $c(t)$, we need to find the conditional distribution of $Y \mid X+Y=t$ when $\theta=\theta$, that is when $p_1 = p_2 = p$

$$P(Y=y \mid X+Y=t) = \frac{P(Y=y, X=t-y)}{P(X+Y=t)}$$

$$= \frac{P(Y=y)P(X=t-y)}{P(X+Y=t)}$$

$$= \frac{\binom{n_2}{y} p^y (1-p)^{n_2-y} \binom{n_1}{t-y} p^t (1-p)^{n_1-t-y}}{\binom{n_1+n_2}{t} p^t (1-p)^{n_1+n_2-t}}$$

$$= \frac{\binom{n_2}{y}\binom{n_1}{t-y}}{\binom{n_1+n_2}{t}} = \text{hypergeometric}(n_2, n_1, t)$$

for $y = 0, 1, 2, \ldots t$

when $p_1 \neq p_2$, it is easily shown that

$$\boxed{P(Y=y \mid X+Y=t) = C_t(\rho) \binom{n_2}{y}\binom{n_1}{t-y} \rho^y} \quad (1)$$

$y = 0, 1, \ldots t$, where

$$C_\rho(t) = \frac{1}{\sum_{j=0}^{t} \binom{n_2}{j}\binom{n_1}{t-j}\rho^j}$$

$\rho = e^\theta$.

Thus

$$E_{\theta_0}[\phi(u) \mid T=t]$$

$$\alpha = P(Y < c(t) \mid T=t) + \gamma(t)\, P(Y = c(t) \mid T=t)$$

$$= \left\{ \sum_{y=0}^{c(t)-1} \frac{\binom{n_2}{y}\binom{n_1}{t-y}}{\binom{n_1+n_2}{t}} \right\} + \gamma(t) \frac{\binom{n_2}{c(t)}\binom{n_1}{t-c(t)}}{\binom{n_1+n_2}{t}}$$

where $c(t)$ is a positive integer.

$$\text{p-value} = P(Y \geq y_{obs} \mid T=t)$$

$$= \sum_{y=y_{obs}}^{t} \frac{\binom{n_2}{y}\binom{n_1}{t-y}}{\binom{n_1+n_2}{t}}$$

(d)

$$\pi(\rho) = P(Y < c(t) \mid T=t) + \gamma(t)\, P(Y = c(t) \mid T=t)$$

$$= \sum_{y=0}^{c(t)-1} C_t(\rho) \binom{n_2}{y}\binom{n_1}{t-y} \rho^y + \gamma(t)\, C_t(\rho) \binom{n_2}{c(t)}\binom{n_1}{t-c(t)} \rho^{c(t)}$$

⋮

e) $n_1 = n_2 = 2$, $X = Y = 1$

$H_0: P_1 \le P_2$
$H_1: P_1 > P_2$

$$B = \frac{\int_0^1 \int_0^{P_2} P_1(1-P_1)\, P_2(1-P_2)\, 2 \, dP_1 dP_2}{\int_0^1 \int_{P_2}^1 P_1(1-P_1)\, P_2(1-P_2)\, 2 \, dP_1 dP_2}$$

under H_0, $\pi(P_1, P_2) = c\, I(P_1 \le P_2)$

$$\int_0^1 \int_0^{P_2} c \, dP_1 dP_2 = 1$$

$$\Rightarrow c = 2$$

under H_1, $\int_0^1 \int_{P_2}^1 c \, dP_1 dP_2 = 1 \Rightarrow c = 2$

under H_0, $\pi(P_1, P_2) = \begin{cases} 2, & P_1 \le P_2 \\ 0 & \text{otherwise} \end{cases}$

under H_1, $\pi(P_1, P_2) = \begin{cases} 2 & P_1 > P_2 \\ 0 & \text{otherwise} \end{cases}$

Thus

$$\text{numerator} = \int_0^1 P_2(1-P_2)\left[\int_0^{P_2} P_1(1-P_1)\,dP_1\right] dP_2$$

$$= \int_0^1 P_2(1-P_2)\left[\frac{P_1^2}{2} - \frac{P_1^3}{3}\Big|_0^{P_2}\right] dP_2$$

$$= \int_0^1 P_2(1-P_2)\left[\frac{P_2^2}{2} - \frac{P_2^3}{3}\right] dP_2$$

$$= \int_0^1 \frac{P_2^3}{2} - \frac{P_2^4}{2} - \frac{P_2^4}{3} + \frac{P_2^6}{3}\,dP_2$$

$$= \frac{P_2^4}{8} - \frac{P_2^5}{10} - \frac{P_2^5}{15} + \frac{P_2^7}{21}\Big|_0^1$$

$$= 1/8 - \frac{1}{10} - \frac{1}{15} + \frac{1}{21} = .005952381$$

$$\text{denominator} = \int_0^1 P_2(1-P_2)\left[\int_{P_2}^1 P_1(1-P_1)\,dP_1\right] dP_2$$

$$= \int_0^1 P_2(1-P_2)\left[\frac{1}{6} - \frac{P_2^2}{2} + \frac{P_2^3}{3}\right] dP_2$$

$$= \int_0^1 \left(\frac{P_2}{6} - \frac{P_2^3}{2} + \frac{P_2^4}{3}\right)(1-P_2)\, dP_2$$

$$= \int_0^1 \frac{P_2}{6} - \frac{P_2^3}{2} + \frac{P_2^4}{3} - \frac{P_2^2}{6} + \frac{P_2^4}{2} - \frac{P_2^5}{3}\, dP_2$$

$$= \frac{P_2^2}{12} - \frac{P_2^4}{8} + \frac{P_2^5}{15} - \frac{P_2^3}{18} + \frac{P_2^5}{10} - \frac{P_2^6}{18}\Bigg|_0^1$$

$$= \frac{1}{12} - \frac{1}{8} + \frac{1}{15} - \frac{1}{18} + \frac{1}{10} - \frac{1}{18} = .0138889$$

$$\frac{.005952381}{.0138889} = .4285714$$

Thus $B = 2.33$ in favor of H_1 (weak evidence in favor of H_1)

Q5 Solutions

(a) i) $$\sum_{n=-\infty}^{\infty} P_n(t) z^n = \sum_{n=-\infty}^{\infty} P(X(t)-Y(t)=n)\, z^n$$

$$= \sum_{n=-\infty}^{\infty} \sum_{x=0}^{\infty} P(X(t)=x, Y(t)=x-n)\, z^n$$

$$= \sum_{n=-\infty}^{\infty} \sum_{x=n}^{\infty} P(X(t)=x, Y(t)=x-n)\, z^n$$

$$= \sum_{x=0}^{\infty} \sum_{n=-\infty}^{x} \frac{(\lambda_1 t)^x e^{-\lambda_1 t}}{x!} \frac{(\lambda_2 t)^{x-n}}{(x-n)!} e^{-\lambda_2 t} z^n$$

$$= \sum_{x=0}^{\infty} \left\{ \frac{(\lambda_1 t)^x e^{-\lambda_1 t}}{x!} (\lambda_2 t)^x e^{-\lambda_2 t} \right\} \sum_{n=-\infty}^{x} \frac{\left(\frac{z}{\lambda_2 t}\right)^n}{(x-n)!}$$

Let $j = x-n$, then the inner sum becomes

$$\sum_{j=0}^{\infty} \frac{(z/\lambda_2 t)^{x-j}}{j!} = \left(\frac{z}{\lambda_2 t}\right)^x \sum_{j=0}^{\infty} \frac{\left(\frac{\lambda_2 t}{z}\right)^j}{j!}$$

$$= \left(\frac{z}{\lambda_2 t}\right)^x e^{\lambda_2 t / z}$$

Now the outer sum becomes

$$= e^{\lambda_2 t/z} e^{-(\lambda_1+\lambda_2)t} \sum_{x=0}^{\infty} \left(\frac{z}{\lambda_2 t}\right)^x \frac{(\lambda_1 t)^x}{x!} (\lambda_2 t)^x$$

$$= e^{\lambda_2 t/z} e^{-(\lambda_1+\lambda_2)t} \sum_{x=0}^{\infty} \frac{(z\lambda_1 t)^x}{x!}$$

$$= e^{\lambda_2 t/z} e^{-(\lambda_1+\lambda_2)t} e^{z\lambda_1 t}$$

$$= e^{-(\lambda_1+\lambda_2)t} e^{\lambda_1 z t + \lambda_2 t/z}$$

ii) Part (i) gives the pgf, so that

if

$$\phi(z) = \sum_{n=-\infty}^{\infty} P_n(t) z^n$$

$$\phi'(1) = E(Z(t))$$

$$\phi''(1) = E(Z(t)(Z(t)-1))$$

$$= E(Z^2(t)) - E(Z(t))$$

Thus $\mathrm{Var}(Z(t)) = \phi''(1) + \phi'(1) - \phi'(1)^2$

(P.3)

Thus

$$\phi'(z) = e^{-(\lambda_1+\lambda_2)t} \, e^{\lambda_1 z t + \frac{\lambda_2 t}{z}} \left[\lambda_1 t - \frac{\lambda_2 t}{z^2}\right]$$

$$\phi'(1) = e^{-(\lambda_1+\lambda_2)t} \, e^{\lambda_1 t + \lambda_2 t} \left[\lambda_1 t - \lambda_2 t\right]$$

$$= (\lambda_1 - \lambda_2)t = E(Z(t))$$

$$\phi''(z) = e^{-(\lambda_1+\lambda_2)t} \, e^{\lambda_1 z t + \frac{\lambda_2 t}{z}} \left[\lambda_1 t - \frac{\lambda_2 t}{z^2}\right]^2$$
$$+ e^{-(\lambda_1+\lambda_2)t} \, e^{\lambda_1 z t + \frac{\lambda_2 t}{z}} \left[2\lambda_2 t z^{-3}\right]$$

$$\phi''(1) = (\lambda_1 t - \lambda_2 t)^2 + 2\lambda_2 t$$

$$\mathrm{Var}(Z(t)) = t^2(\lambda_1-\lambda_2)^2 + 2\lambda_2 t + (\lambda_1-\lambda_2)t - (\lambda_1-\lambda_2)^2 t^2$$

$$= 2\lambda_2 t + \lambda_1 t - \lambda_2 t$$

$$= (\lambda_1+\lambda_2)t$$

(b)

$$P\left(Y(t) = x - x_0 \mid X(t) + Y(t) = z - x_0 - y_0\right)$$

$$= \frac{P\left(X(t) = x - x_0,\ Y(t) = z - y_0 - x\right)}{P\left(X(t) + Y(t) = z - x_0 - y_0\right)}$$

$$= \frac{P\left(X(t) = x - x_0\right) P\left(Y(t) = z - y_0 - x\right)}{P\left(X(t) + Y(t) = z - x_0 - y_0\right)}$$

$$= \frac{\dfrac{(\lambda_1 t)^{x - x_0} e^{-\lambda_1 t}}{(x - x_0)!} \cdot \dfrac{(\lambda_2 t)^{z - y_0 - x} e^{-\lambda_2 t}}{(z - y_0 - x)!}}{\dfrac{\left((\lambda_1 + \lambda_2) t\right)^{z - x_0 - y_0} e^{-(\lambda_1 + \lambda_2) t}}{(z - x_0 - y_0)!}}$$

or substitute $z = x + y$

$$= \binom{x + y - x_0 - y_0}{x - x_0} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{x - x_0} \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{y - y_0}, \quad x \ge x_0,\ y \ge y_0$$

(c)

i) $P(Z_1 > t) = P(X(t) = 0) = e^{-\lambda_1 t}$

$P(Z_2 > t \mid Y_1 = s) = P(0 \text{ events in } (s, s+t] \mid Z_1 = s)$

$= P(0 \text{ events in } (s, s+t])$ by indep increments

$= e^{-\lambda_1 t}$

Thus $Z_1 .. Z_n$ are iid exponential λ_1. We want the density of

$$S_n = \sum_{i=1}^{n} Z_i = \sum_{i=1}^{n} \text{iid } exp(\lambda_1) = gamma(n, \lambda_1)$$

We can derive this using MGF's or by noting

$$F_{S_n}(t) = P(S_n \le t) = P(X(t) \ge n)$$

$$= \sum_{j=n}^{\infty} e^{-\lambda_1 t} \frac{(\lambda_1 t)^j}{j!}$$

$$f_{S_n}(t) = \frac{d}{dt} F_{S_n}(t) = -\sum_{j=n}^{\infty} \lambda_1 e^{-\lambda_1 t} \frac{(\lambda_1 t)^j}{j!} + \sum_{j=n}^{\infty} \frac{\lambda_1 e^{-\lambda_1 t} (\lambda_1 t)^{j-1}}{(j-1)!}$$

$$= \frac{\lambda_1 e^{-\lambda_1 t} (\lambda_1 t)^{n-1}}{(n-1)!} = gamma(n, \lambda_1).$$

(ii) $f(S_1, \dots S_n \mid X(t)=n)$

$= P(t_i \le S_i \le t_i + h_i,\ i=1,\dots,n \mid X(t)=n)$

$$= \frac{P(\text{exactly } 1 \text{ event in } [t_i, t_i+h_i],\ i=1,\dots,n,\ \text{no events elsewhere in } [0,t))}{P(X(t)=n)}$$

$$= \frac{(\lambda_1 h_1 e^{-\lambda_1 h_1})(\lambda_1 h_2 e^{-\lambda_1 h_2})\cdots(e^{-\lambda_1(t-h_1-h_2-\dots-h_n)})}{\frac{e^{-\lambda_1 t}(\lambda_1 t)^n}{n!}}$$

$$= \frac{n!}{t^n} h_1 \cdots h_n$$

Thus

$$\frac{P(t_i \le S_i \le t_i+h_i,\ i=1,\dots,n \mid X(t)=n)}{h_1 \cdots h_n} = \frac{n!}{t^n}$$

Letting $h_i \to 0$, we get

$$f(S_1, \ldots, S_n \mid X(t)=n) = \frac{n!}{t^n}, \quad 0 < t_1 < \cdots < t_n.$$

I) We start out with i individuals

$$P_{in}(t) = P(X(t)=n \mid X(0)=i) = P(n \text{ people alive at time } t \mid X(0)=i)$$

$$= P(i-n \text{ deaths} \mid X(0)=i)$$

Since deaths occur exponentially, the

P(survival for an individual)

$$= P(T>t) = e^{-\mu t}$$

$$P(\text{death}) = 1-P(T\leq t) = 1-e^{-\mu t}$$

Thus from the i initial persons any $i-n$ of them can die with prob. $1-e^{-\mu t}$

There are $\binom{i}{i-n}$ ways to choose the dead persons

Thus

$$P_{in}(t) = \binom{i}{i-n}\left(1-e^{-\mu t}\right)^{i-n}\left(e^{-\mu t}\right)^{n}$$

Thus

$$E(X(t)) = i\,e^{-\mu t}, \quad \mathrm{Var}(X(t)) = i\,e^{-\mu t}\left(1-e^{-\mu t}\right)$$