

# STAT 301: Homework set 7

Daniel Li

## Problem 1

### Subproblem 1

$$\mu \exp(-\mu x) = \exp(-\mu x + \log(\mu))$$

$$h(x) = 1$$

$$t(x) = x$$

$$\eta = -\mu$$

$$a(\eta) = -\log(-\eta)$$

$$\frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx) = \exp(a \log(b) + (a-1) \log(x) - bx - \log(\Gamma(a)))$$

$$h(x) = 1$$

$$t(x) = \begin{bmatrix} \log(x) \\ x \end{bmatrix}$$

$$\eta = \begin{bmatrix} a-1 \\ -b \end{bmatrix}$$

$$a(\eta) = \log(\Gamma(\eta_1 + 1)) - (\eta_1 + 1) \log(-\eta_2)$$

### Subproblem 2

$$P(\mu | a, b) = \exp(a \log(b) + (a-1) \log(\mu) - b\mu - \log(\Gamma(a))) = \exp(-b\mu + (a-1) \log(\mu) + a \log(b) - \log(\Gamma(a)))$$

and note that multiplying by  $P(x | \mu)$  preserves the form as only the  $\mu$  terms are affected. Thus, the posterior is also a gamma distribution with parameters  $a' = a + n$  and  $b' = b + \sum_{i=1}^n x_i$ .

### Subproblem 3

$$P(\mu | x_{1:n}, a, b) \propto \exp(-b\mu + (a-1) \log(\mu) + a \log(b) - \log(\Gamma(a))) \prod_{i=1}^n \exp(-\mu x_i + \log(\mu))$$

$$\propto \exp(-\mu(b + \sum_{i=1}^n x_i) + (a+n-1) \log(\mu) + (a \log(b) - \log(\Gamma(a))))$$

$$P(\mu | x_{1:n}, a, b) = \exp(-\mu(b + \sum_{i=1}^n x_i) + (a+n-1) \log(\mu) + ((a+n) \log(b + \sum_{i=1}^n x_i) - \log(\Gamma(a+n))))$$

### Subproblem 4

$$\begin{aligned}
P(x_1 | a, b) &= \int P(x_1 | \mu) P(\mu | a, b) d\mu \\
&= \int \exp(-\mu x_1 + \log(\mu)) \exp(-b\mu + (a-1)\log(\mu) + a\log(b) - \log(\Gamma(a))) d\mu \\
&= \exp(a\log(b) - \log(\Gamma(a))) \int \exp(-\mu(b+x_1) + a\log(\mu)) d\mu \\
&= \exp(a\log(b) - \log(\Gamma(a))) \int \exp(-\mu(b+x_1) + a\log(\mu)) d\mu \\
&= \frac{b^a}{\Gamma(a)} \int_0^\infty \mu^a \exp(-\mu(b+x_1)) d\mu \\
&= \frac{b^a}{\Gamma(a)} \frac{\Gamma(a+1)}{(x_1+b)^{a+1}} \\
&= \frac{b^a a}{(x_1+b)^{a+1}}
\end{aligned}$$

### Subproblem 5

$$\begin{aligned}
P(x_{n+1} | x_{1:n}, a, b) &= \frac{(b + \sum_{i=1}^n x_i)^{a+n}}{\Gamma(a+n)} \int \mu^{a+n} \exp(-\mu(x_{n+1} + b + \sum_{i=1}^n x_i)) d\mu \\
&= \frac{(b + \sum_{i=1}^n x_i)^{a+n}}{\Gamma(a+n)} \frac{\Gamma(a+n+1)}{(x_{n+1} + b + \sum_{i=1}^n x_i)^{a+n+1}} \\
&= \frac{(b + \sum_{i=1}^n x_i)^{a+n} (a+n)}{(x_{n+1} + b + \sum_{i=1}^n x_i)^{a+n+1}}
\end{aligned}$$

## Problem 2

### Subproblem 1

$$\begin{aligned}
P(\mu | x, a, b) &\propto \frac{\mu^x}{x!} \exp(-\mu) \exp(a\log(b) + (a-1)\log(\mu) - b\mu - \log(\Gamma(a))) \\
&\propto \mu^{x+a-1} \exp(-\mu(b+1)) \\
P(\mu | x, a, b) &= \frac{\mu^{x+a-1} \exp(-\mu(b+1)) (b+1)^{a+x}}{\Gamma(x+a)}
\end{aligned}$$

### Subproblem 2

$$\begin{aligned}
\frac{\mu^x}{x!} \exp(-\mu) &= \frac{1}{x!} \exp(\log(\mu)x - \mu) \\
h(x) &= \frac{1}{x!} \\
t(x) &= x \\
\eta &= \log(\mu) \\
a(\eta) &= \exp(\eta)
\end{aligned}$$

$$\begin{aligned}
\frac{\Gamma(x+r)}{x!\Gamma(r)}(1-p)^r p^x &= \frac{\Gamma(x+r)}{x!\Gamma(r)} \exp(\log(p)x + r \log(1-p)) \\
h(x) &= \frac{\Gamma(x+r)}{x!\Gamma(r)} \\
t(x) &= x \\
\eta &= \log(p) \\
a(\eta) &= -r \log(1 - \exp(\eta))
\end{aligned}$$

### Subproblem 3

$$\begin{aligned}
&\sum_{x \in \mathbb{N}} \frac{1}{x!} \exp(\log(\mu_1)x - \mu_1) \log \left( \frac{\frac{1}{x!} \exp(\log(\mu_1)x - \mu_1)}{\frac{1}{x!} \exp(\log(\mu_2)x - \mu_2)} \right) \\
&= \sum_{x \in \mathbb{N}} \frac{1}{x!} \exp(\log(\mu_1)x - \mu_1) (x \log(\frac{\mu_1}{\mu_2}) - \mu_1 + \mu_2) \\
&= \mu_1 \log(\frac{\mu_1}{\mu_2}) - \mu_1 + \mu_2
\end{aligned}$$

### Subproblem 4

$$\begin{aligned}
H(Pois(y) | Gam(a, b)) &= \int_0^\infty - \sum_x \frac{y^x}{x!} e^{-y} \log\left(\frac{y^x}{x!} e^{-y}\right) dGam(a, b) \\
&= \int_0^\infty - \sum_x \frac{y^x}{x!} e^{-y} \log\left(\frac{y^x}{x!} e^{-y}\right) dGam(a, b) \\
&\leq - \sum_x NB(x; a, \frac{1}{1+b}) \log(NB(x; a, \frac{1}{1+b})) \\
&= H(NB(x; a, \frac{1}{1+b}))
\end{aligned}$$

Where we use Jensen's inequality on  $f(t) = t \log(t)$ .

## Problem 3

### Subproblem 1

$$\begin{aligned}
\operatorname{argmax}_k H(Z) - H(Z | Y_k) &= \operatorname{argmax}_k H(Y_k) - H(Y_k | Z) \\
&= \operatorname{argmax}_k H(Y_k) \\
&= \operatorname{argmax}_k H_2(\pi_k) \\
&= \operatorname{argmax}_k \pi_k
\end{aligned}$$

This works trivially when  $K \leq 2$  although the  $K = 2$  case is more of an arbitrary choice. Otherwise, the last line follows because the binary entropy is symmetrical about 0.5 so we have the cases 1) where all

$\pi_k < 0.5$  then the maximum is closest to 0.5 and 2) where there is one  $\pi_j > 0.5$  and that is closest to 0.5 because every other  $\pi_k$  is less than  $1 - \pi_j$ .

### Subproblem 2

$$\begin{aligned} \operatorname{argmax}_k H(Z) - H(Z | Y_k) &= \operatorname{argmax}_k H(Y_k) - H(Y_k | Z) \\ &= \operatorname{argmax}_k [-\pi_k q_k \log(\pi_k q_k) - (1 - \pi_k q_k) \log(1 - \pi_k q_k)] + \pi_k (q_k \log(q_k) + (1 - q_k) \log(1 - q_k)) \\ &= \operatorname{argmax}_k H_2(\pi_k q_k) + \pi_k H_2(q_k) \end{aligned}$$

### Subproblem 3

pi	0.75	0.125	0.06125	0.06125
q	0.25	0.5	0.75	0.5
H(Y)	0.69621226	0.33729007	0.26888241	0.19751658
	0.81127812	1	0.81127812	1
$H(Y Z)$	0.60845859	0.125	0.04969079	0.06125
H(Y) - $H(Y Z)$	0.08775367	0.21229007	0.21919163	0.13626658

$k^* = 2$ . It is not what I expected because of the low pi. But it makes sense in that there is more information in finding the submarine in a cell that you didn't expect it from. The relationship with SEPs and optimal search is that higher SEP means that the search is more likely to be successful so you would reduce your uncertainty by searching in the cells with higher SEP.