Josh Kutta
Prof. Paisley
ELEN4903
HW 1.

Problem 1

a) $p(x_1, x_2, \ldots, x_N \mid \pi) = \prod_{i=1}^{N} p(x_i \mid \pi) = \prod_{i=1}^{N} \pi^{x_i}(1-\pi)^{1-x_i} = \pi^{\sum_{i=1}^N x_i}(1-\pi)^{\sum_{i=1}^N (1-x_i)} = \pi^{\sum_{i=1}^N x_i}(1-\pi)^{N-\sum_{i=1}^N x_i}$

b) $\hat{\pi}_{ML} = \arg\max_{\pi} \prod_{i=1}^{N} p(x_i \mid \pi) \Rightarrow \nabla_\pi \ln\left(\pi^{\sum_{i=1}^N x_i}(1-\pi)^{N-\sum_{i=1}^N x_i}\right) = 0$

$\Rightarrow \nabla_\pi \left(\sum_{i=1}^N x_i\right)\ln(\pi) + \left(N - \sum_{i=1}^N x_i\right)\ln(1-\pi) = 0, \quad \frac{1}{\pi}\sum_{i=1}^N x_i - \frac{1}{1-\pi}\left(N - \sum_{i=1}^N x_i\right) = 0,$

$\Rightarrow \frac{1-\pi}{\pi} = \frac{N - \sum_{i=1}^N x_i}{\sum_{i=1}^N x_i}, \quad \frac{1}{\pi} - \cancel{1} = \frac{N}{\sum_{i=1}^N x_i} - \cancel{1}, \quad \boxed{\hat{\pi}_{ML} = \frac{\sum_{i=1}^N x_i}{N}}$

(c) $\hat{\pi}_{MAP} = \arg\max_\pi \ln p(\pi \mid x_1, \ldots, x_N) \Rightarrow \nabla_\pi \ln\left(\pi^{\sum_{i=1}^N x_i}(1-\pi)^{N-\sum_{i=1}^N x_i}\right) + \ln(\underbrace{Beta(\pi \mid a,b)}_{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\pi^{a-1}(1-\pi)^{b-1}}) = 0$

$\Rightarrow \underbrace{\left(\frac{1}{\pi}\sum_{i=1}^N x_i - \frac{1}{1-\pi}\left(N - \sum_{i=1}^N x_i\right)\right)}_{\text{from (b)}} + \nabla_\pi \left(\cancel{\ln\left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right)}^{0} + (a-1)\ln(\pi) + (b-1)\ln(1-\pi)\right) = 0$

$\Rightarrow \left(\frac{1}{\pi}\sum_{i=1}^N x_i - \frac{1}{1-\pi}\left(N - \sum_{i=1}^N x_i\right) + \frac{a-1}{\pi} - \frac{b-1}{1-\pi} = 0\right)\pi(1-\pi)$

$\Rightarrow (1-\pi)\sum_{i=1}^N x_i - \pi\left(N - \sum_{i=1}^N x_i\right) + (1-\pi)(a-1) - \pi(b-1) = 0$

$\Rightarrow (1-\pi)\left(\sum_{i=1}^N (x_i) + a - 1\right) = \pi\left(N - \sum_{i=1}^N (x_i) + b - 1\right) \Rightarrow \sum_{i=1}^N (x_i) + a - 1 = \pi(N + a + b - 2)$

$\Rightarrow \boxed{\hat{\pi}_{MAP} = \frac{\sum_{i=1}^N x_i + a - 1}{N + a + b - 2}}$

(d) $p(\pi \mid x_1, x_2, \ldots, x_N) = \frac{p(x_1, \ldots, x_N \mid \pi)\, p(\pi)}{p(x_1, \ldots, x_N)} = \frac{p(x_1, \ldots, x_N \mid \pi)\, p(\pi)}{\int_0^1 p(x_1, \ldots, x_N \mid \pi)\, p(\pi)\, d\pi}$ ← constant (indep. of π)

$p(\pi \mid x_1, \ldots, x_N) \propto p(x_1, \ldots, x_N \mid \pi)\, p(\pi) = \left(\pi^{\sum_{i=1}^N x_i}(1-\pi)^{\sum_{i=1}^N (1-x_i)}\right)\cdot\left(\underbrace{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}}_{\text{constant}}\pi^{a-1}(1-\pi)^{b-1}\right)$

$p(\pi \mid x_1, \ldots, x_N) \propto \pi^{\sum_{i=1}^N (x_i) + a - 1}(1-\pi)^{\sum_{i=1}^N (1-x_i) + b - 1}$

Since we know $\int_0^1 Beta(\pi \mid a,b)\, d\pi = 1 \Rightarrow p(\pi \mid x_1, \ldots, x_N) = \frac{\Gamma(N+a+b)}{\Gamma\left(\sum_{i=1}^N x_i + a\right)\Gamma\left(\sum_{i=1}^N (1-x_i) + b\right)}\pi^{\sum_{i=1}^N (x_i) + a - 1}(1-\pi)^{\sum_{i=1}^N (1-x_i) + b - 1}$

$p(\pi \mid x_1, \ldots, x_N) = Beta\left(\sum_{i=1}^N x_i + a, \sum_{i=1}^N (1-x_i) + b\right)$ (Beta distribution)

(e)(i) set $a' = \sum_{i=1}^N x_i + a$, $b' = \sum_{i=1}^N (1-x_i) + b$; $\mathbb{E}[\pi] = \int_0^1 \pi \frac{\Gamma(a'+b')}{\Gamma(a')\Gamma(b')}\pi^{a'-1}(1-\pi)^{b'-1}\, d\pi$

$\int_0^1 \frac{\Gamma(a'+b')}{\Gamma(a')\Gamma(b')}\pi^{a'-1}(1-\pi)^{b'-1}\, d\pi = 1 \Rightarrow \frac{\Gamma(a')\Gamma(b')}{\Gamma(a'+b')} = \int_0^1 \pi^{a'-1}(1-\pi)^{b'-1}\, d\pi \Rightarrow \frac{\Gamma(a'+1)\Gamma(b')}{\Gamma(a'+b'+1)} = \int_0^1 \pi^{a'}(1-\pi)^{b'-1}\, d\pi$

$\Rightarrow \mathbb{E}[\pi] = \frac{\Gamma(a'+b')}{\Gamma(a')\Gamma(b')}\int_0^1 \pi^{a'}(1-\pi)^{b'-1}\, d\pi = \frac{\Gamma(a'+b')\Gamma(a'+1)\cancel{\Gamma(b')}}{\Gamma(a')\cancel{\Gamma(b')}\Gamma(a'+b'+1)} = \frac{\cancel{\Gamma(a'+b')}\, a'\, \cancel{\Gamma(a')}}{\cancel{\Gamma(a')}(a'+b')\cancel{\Gamma(a'+b')}} = \frac{a'}{a'+b'}$

$\mathbb{E}[\pi] = \frac{\sum_{i=1}^N x_i + a}{N + a + b}$

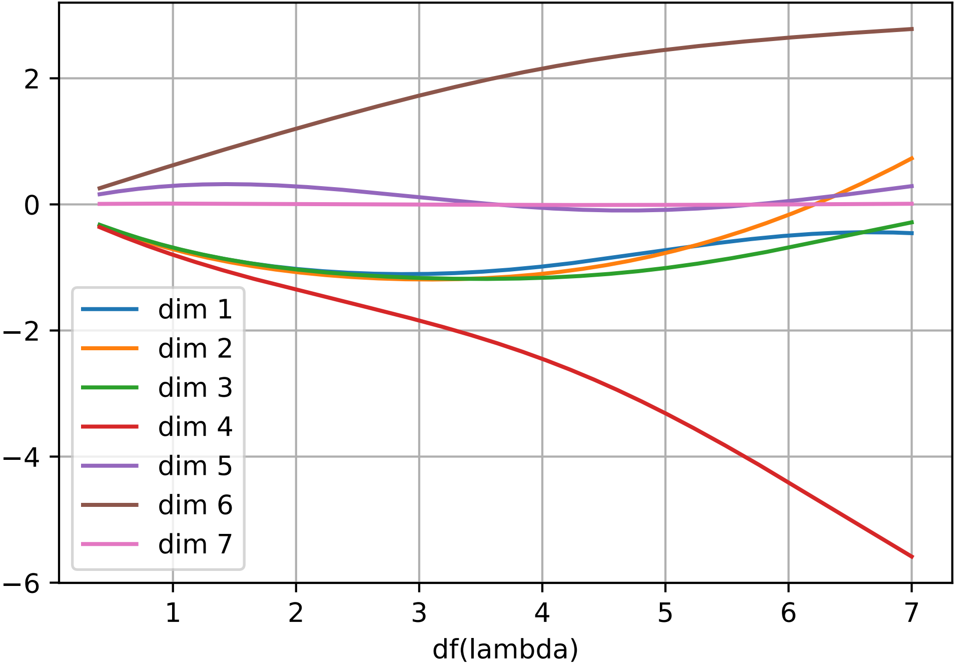
(ii) $Var(\pi) = \mathbb{E}[\pi^2] - \mathbb{E}[\pi]^2$; $\mathbb{E}[\pi^2] = \frac{\Gamma(a'+b')}{\Gamma(a')\Gamma(b')}\int_0^1 \pi^{a'+1}(1-\pi)^{b'-1}\, d\pi = \frac{\Gamma(a'+b')}{\Gamma(a')\cancel{\Gamma(b')}}\cdot\frac{\Gamma(a'+2)\cancel{\Gamma(b')}}{\Gamma(a'+b'+2)} = \frac{(a'+1)a'}{(a'+b'+1)(a'+b')}$

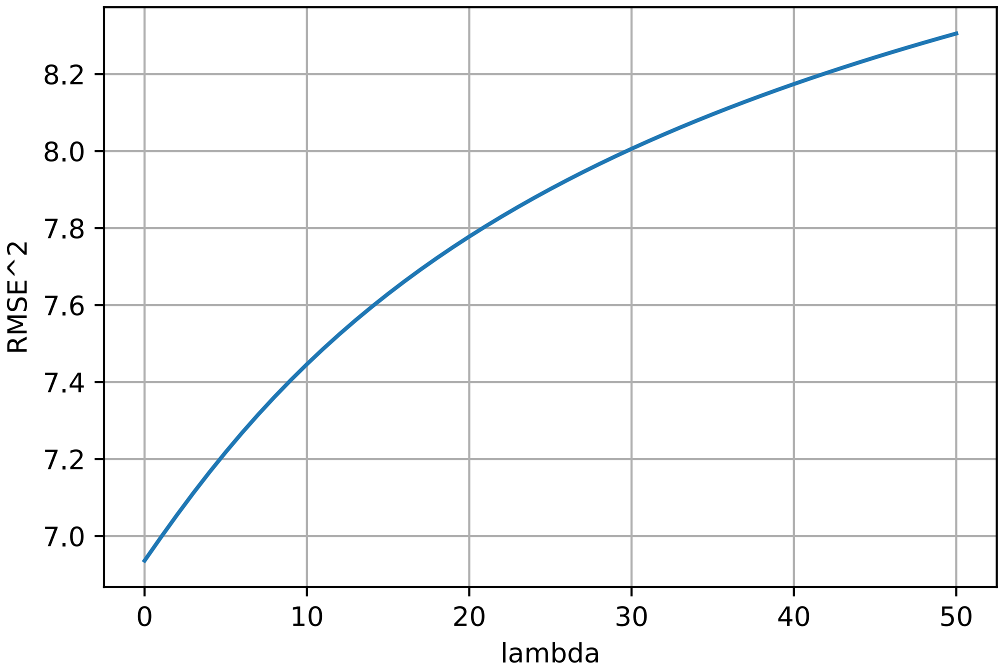
1. (e) (ii) (cont'd) $\mathbb{E}[\pi^2] = \frac{(a'+1)a'}{(a'+b'+1)(a'+b')}$, $\mathbb{E}[\pi]^2 = \frac{(a')^2}{(a'+b')^2}$

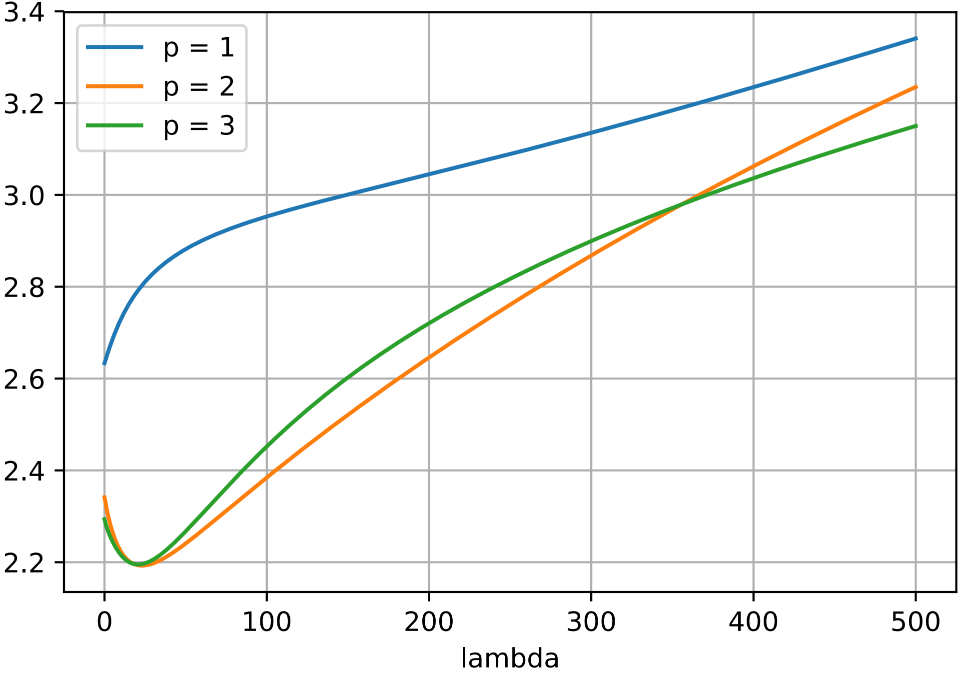
$$\mathrm{var}(\pi) = \mathbb{E}[\pi^2] - \mathbb{E}[\pi]^2 = \frac{(a'^2+a')(a'+b')}{(a'+b'+1)(a'+b')^2} - \frac{(a')^2(a'+b'+1)}{(a'+b'+1)(a'+b')^2} = \frac{\cancel{a'^3} + \cancel{a'^2(b'+1)} + a'b' - [\cancel{a'^3} + \cancel{a'^2(b'+1)}]}{(a'+b'+1)(a'+b')^2}$$

$$\mathrm{var}(\pi) = \frac{a'b'}{(a'+b'+1)(a'+b')^2}$$

$$\mathrm{Var}(\pi) = \frac{\left(\sum_{i=1}^{N} x_i + a\right)\left(\sum_{i=1}^{N}(1-x_i) + b\right)}{(N+a+b+1)(N+a+b)}$$


2
0
−2
−4
−6
1
2
3
4
5
6
7
df(lambda)
dim 1
dim 2
dim 3
dim 4
dim 5
dim 6
dim 7


8.2
8.0
7.8
7.6
7.4
7.2
7.0
RMSE^2
0
10
20
30
40
50
lambda


p = 1
p = 2
p = 3
3.4
3.2
3.0
2.8
2.6
2.4
2.2
0
100
200
300
400
500
lambda