PSIS + L-O-O

s \ i	1	2	3	$\prod_{i,s} \Pr(y_i \mid \theta_s)$	$\log(\Pr(y_i \mid \theta_{-j,s}))$	$\log\left(\frac{1}{\Pr(y_i \mid \theta_{-j})}\right) = r_{-j}$ (Importance)	$\log\left(\frac{r_{-j}\Pr(y_i \mid \theta_{-j})}{R}\right)$
	$\Pr(y_i \mid \theta_s)$: $1/2$	$1/3$	$1/6$				
0	[redacted]	[redacted]	[redacted]	$(\frac{1}{2})(\frac{1}{3})(\frac{1}{6}) = \frac{1}{36}$	$\log(\frac{1}{36}) = -3.58$		
1	O	[redacted]	[redacted]	$(\frac{1}{3})(\frac{1}{6}) = \frac{1}{18}$	$\log(\frac{1}{18}) = -2.89$	$\log(\frac{1}{1/2}) = \log(2) = 0.69$	$\log(0.69(1/18)/3.57) = -4.53$
2	[redacted]	O	[redacted]	$(1/2)(1/6) = \frac{1}{12}$	$\log(\frac{1}{12}) = -2.49$	$\log(\frac{1}{1/3}) = \log(3) = 1.09$	$\log(1.09(1/12)/3.57) = -3.67$
3	[redacted]	[redacted]	O	$(1/2)(1/3) = 1/6$	$\log(1/6) = -1.79$	$\log(\frac{1}{1/6}) = \log(6) = 1.79$ 50%!	$\log(1.79(1/6)/3.57) = -2.48$

$$R = \sum_s \left(\frac{1}{\Pr(y_i \mid \theta_{-j})}\right) = 3.57$$

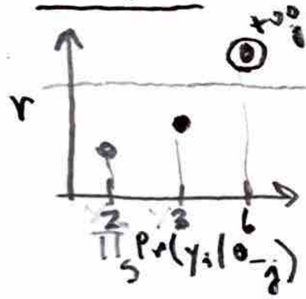
$$S = \text{lppd}_{cv} = \log(\exp(-2.89) + \exp(-2.49) + \exp(-1.79)) - \log(3) = -1.09$$

$$\boxed{\text{lppd}_{cv} = -1.09}$$

$$\boxed{\text{lppd}_{IS} = -3.31}$$

$\log(\exp(-4.53) + \exp(-3.67) + \exp(-2.48)) = \log(0.02 + 0.03 + 0.06) = \log(0.11) - \log(3) = -3.31$

PS part



$\bar{u} = 1.19$

$\sigma = 0.75$

s	Impr.	k-Smoothed	WGT		log	exp
1	0.69	0.69	0.33	$\log((0.33)(1/18)) =$	-4.00	0.02
2	1.09	1.09	0.52	$\log((0.52)(1/12)) =$	-3.14	0.04
3	(0.169)(1.79)	0.30	0.15	$\log((0.15)(1/6)) =$	-3.69	0.03
		Σ 2.08				

$\log(0.11) - \log(3) = -3.31$

$\log(0.09/3) = -3.62$

$$\boxed{\text{PSIS} = -3.62}$$

Max Likelihood est. $k = 0.7$		
k	$-\frac{1}{k} - 1$	$\Pr(r \mid k)$
0.1	-11.0	0.019
0.3	-4.3	0.116
0.5	-3.0	0.154
0.7	-2.43	0.169

$$r \sim \frac{1}{\sigma}\left(1 + \frac{k}{\sigma}(r - u)\right)^{-\frac{1}{k} - 1} = (1.33 + 1.07k)^{-\frac{1}{k} - 1}$$

GPD(u = loc, σ = scale, k = shape)

$u = \bar{r}$	$\sigma = s_r$
$= 1.19$	$= 0.75$

$(0.69 - 1.19)^2 = 0.25$

$(1.09 - 1.19)^2 = 0.01$

$(1.79 - 1.19)^2 = 1.41$

$1.67/3 = 0.56$

PSIS/WAIC 1/2

! Simulate !

WE GET TO WAIC

$$\text{WAIC} = -2\,(\text{lppd} - p_{\text{WAIC}})$$

AS (lppd)

"effective parameters"

"Resolves Singularities"

Allows us to cross over

$$p_{\text{WAIC}} = \text{Var}(\log(\Pr(y_i \mid \theta)))$$

H(p)

A

$\frac{1}{2}\log(\frac{1}{2}) = \frac{1}{2}(2.89) = 1.45$

$\frac{1}{3}\log(\frac{1}{3}) = \frac{1}{3}(2.99) = 0.99$

$\frac{1}{6}\log(\frac{1}{6}) = \frac{1}{6}(1.79) = 0.30$

B

→ 2.2/3

→ 2.2/3

→ 2.2/3

$H(p_A) = 1.84$

$H(p_B) = 2.20$

WHY?

NN ML models have parameter spaces which have Singularities

All of our conventional stats require us to Invert a matrix

↳ Fisher Information Matrix

N Network

Gaussian Mix

Bayes Network

etc.

exp(log(Pr))	s	log(Pr)	Var( )
$\frac{1}{18}$	1	2.89	$(2.89-2.56)^2 = 0.11$
$\frac{1}{12}$	2	2.99	$(2.99-2.56)^2 = 0.18$
$\frac{1}{6}$	3	2.79	$(2.79-2.56)^2 = 0.59$

$\Sigma = 6/18 = 1/3$

$\log(\frac{1}{3}) = -1.09 = \text{lppd}_A$

$0.29 = \frac{0.88}{3}$

vs

s	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$		
1	☐	$\frac{1}{3} \times \frac{1}{3}$		$= \frac{1}{9}$	$\log(\frac{1}{9}) = -2.20$
2	$\frac{1}{3}$	☐	$\times \frac{1}{3}$	$= \frac{1}{9}$	$\log(\frac{1}{9}) = -2.20$
3	$\frac{1}{3} \times \frac{1}{3}$		☐	$= \frac{1}{9}$	$\log(\frac{1}{9}) = -2.20$

$-2.2 \Leftarrow -6.60/3$

$$\boxed{\begin{aligned}\text{WAIC}_A &= -2(-1.09 - 0.29)\\ &= 2.18 + 0.58\\ &= 2.76\end{aligned}}$$

Just like H(p)

1	2.2
2	2.2
3	2.2

$p_{\text{WAIC}} = 0$

b/c Var = 0

vs

$$\boxed{\begin{aligned}\text{WAIC}_B &= -2(-2.2 - 0)\\ &= 4.4\ !\end{aligned}}$$

M	WAIC	pWAIC
A	2.76	0.29
B	4.40	0

Sample → parameters → prediction

RATHER

samples → functions → prediction

PSIS/WAIC 2/2