Lecture 17

Priors / Jeffrey's : $P(\sigma, \theta | x) \propto P(x|\theta, \sigma^2) P(\theta, \sigma^2)$

$P(\theta, \sigma^2) = P(\theta|\sigma^2) P(\sigma^2)$ Bayes rule

$$P(\theta, \sigma^2) \propto (1)\left(\frac{1}{\sigma^2}\right)$$

- Random Sampling given parameter x ; distributed along $\mathcal{F}(x) = y$

generally
1. select some y from parameter space "Θ_y"
2. find x from $\mathcal{F}^{-1}(y) = x$ } occurs once with each parameter
3. return sample value

** r "function" (parameters) distri. samples in computer

— ex: 1dimensional / Parameter ; Discrete distribution
- $X \sim \text{Binomial}(n, \theta)$, $n = 6$, $\theta = 0.4$, $\text{supp}[X] = \{1,2,3,4,5,6\}$

$F(X) = \text{CDF}$

$F(0) = p(0)$	$F(1) = p(1) + F(0)$	$F(2) =$	$F(3) =$	$F(4) =$	$F(5) =$	$F(6)$
$= 0.047$	$= 0.233$	$= 0.544$	$= 0.821$	$= 0.959$	$= 0.996$	$= 1.$

Sample :
1) select u from $U(0,1)$ (space of cdf)
2) $F^{-1}(u)$ such that, if u falls in range of $[F(x_0), F(x_1))$ $x = x_0$
3) x_0 is our sample value

— ex 2 : 1-d / parameter ; Continuous distribution
- looking for distribution of $F_x(x)$, probability (of x) = y
- Let $F_x(x) = f(x) = Y$ such that $F_Y(Y) \sim [0,1]$ CDF

so,
- $F_Y(y) = P(Y \leq y)$

def of cdf
$$= P(F_x(x) \leq y)$$
$$= P(F_x^{-1}(F_x(x)) < F_x^{-1}(y))$$
$$= P(X < F_x^{-1}(y))$$

$$\Rightarrow F_x(F_x^{-1}(y)) = y$$

* assumes inverse is known

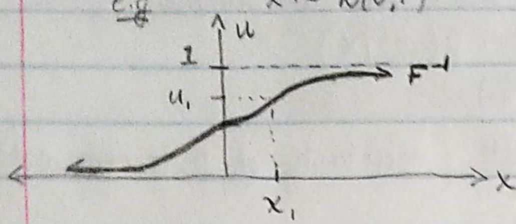
CDF = y } single value y

support : $[0,1]$ } for all of Y

$Y \sim U(0,1)$

$F_x(x) \sim U(0,1)$

e.g. $x \sim N(0,1)$



Sampling

1) draw u from $U(0,1)$
2) let $x = F^{-1}(u)$

* again assuming F^{-1} is Known

- if F^{-1} is unknown : Approximate X r.v. with a Discrete version

$$\underset{\text{grid}}{G} = \langle \underset{x_0}{x_{min}},\ \underset{x_1}{x_{min}+\Delta x},\ \underset{x_2}{x_{min}+2\Delta x},\ \dots,\ \underset{x_n}{x_{max}} \rangle \Big\}\ g \text{ total elements}$$

* Calculate each value by the Riemann Sum of each interval
** then sample from the discrete

— ex 3 : 2-d/parameter ; discrete distrib.

- Given $P(X,Y)$

e.g.

X \ Y	2	4	6	$P(X)$
1	0.2	0.05	0.1	.35
3	0.1	0.05	0.1	.25
5	0.2	0.1	0.1	.40
				1.00

By Bayes rule : $P(X,Y) = P(Y|X)P(X)$
$\boxed{\text{OR} = P(X|Y)\,P(Y)}$

- To sample :

1) Pick x_0
2) Pick y_0 given that x_0
3) Return values $\langle x_0, y_0 \rangle$

— ex 4 : 2-d/parameters ; continuous distrib.

- $X_1 \dots X_n \sim N(\theta, \sigma^2)$ where θ and σ^2 are unknown/variables.

$$P(\theta,\sigma^2|X) = \underbrace{P(\sigma^2|\theta,X)}_{\text{InvGamma}(n/2,\ \frac{n\hat{\sigma}^2}{2})} \cdot P(\theta|X) \quad \leftarrow \text{Most Practical}$$

$$= \underbrace{P(\theta|\sigma^2,X)}_{\text{Norm}\left(\bar{x},\left(\frac{\sigma}{\sqrt{n}}\right)^2\right)} \cdot P(\sigma^2|X)$$

Marginal Posteriors : Integrates out uncertainty from $P(\theta,\sigma^2|X)$

$\left\{ \text{assuming jeffrey's } P(\theta,\sigma^2) \propto \frac{1}{\sigma^2} \right\}$

$\propto \text{NormInvGamma}\left(\bar{x}, n, \frac{n}{2}, \frac{(n-1)s^2}{2}\right)$ Note: s^2 is sample variance

*We must find the Marginal Posteriors before we can sample

Marginal $\sigma^2 | X$

- 1st/ $P(\theta, \sigma^2 | X) = P(\theta | \sigma^2, x) \cdot P(\sigma^2 | X)$
  - $\rightarrow$ NormInvGam$(\mu, \lambda, \alpha, \beta)$ $\rightarrow$ Norm$(\bar{x}, \frac{\sigma^2}{n})$?

- $P(\sigma^2 | X) = \int_{\mathbb{R}} P(\theta, \sigma^2 | X) \, d\theta$ (uncertainty of θ is removed)

- $$P(\sigma^2 | X) = \frac{P(\theta, \sigma^2 | X)}{P(\theta | \sigma^2, x)} \propto \frac{\left(\sigma^{2^{-\frac{n}{2}-1}}\right)\left(e^{\frac{-(n-1)s^2/2}{\sigma^2}}\right)\left(\cancel{e^{\frac{-1}{2\sigma^2}(\bar{x}-\theta)^2}}\right)}{\left(\sigma^{2^{-1/2}}\right)\left(\cancel{e^{\frac{-1}{2\sigma^2}(\bar{x}-\theta)^2}}\right)}$$

Note: $(n-1)s^2 = \Sigma(x_i - \bar{x})^2$

$$\boxed{P(\sigma^2 | X)} = \sigma^{2^{-\left(\frac{n-1}{2}\right)-1}} \, e^{\frac{-(n-1)s^2/2}{\sigma^2}} \propto \boxed{\text{InvGamma}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)}$$

- Hence to Sample from $\theta, \sigma^2 | X \sim$ NormInvGam$(\mu, \lambda, \alpha, \beta)$:

  1) Select σ_0^2 from InvGamma$\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$
  2) select θ_0 from Norm$(\bar{x}, \sigma_0^2/n)$
  3) Return $\langle \theta_0, \sigma_0^2 \rangle$

- Note: $P(\sigma^2 | x) = \text{InvGamma}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$ $\leftarrow$ if $s^2 = \hat{\sigma}^2$, distribution is over-dispersed

  vs

  $P(\sigma^2 | x, \theta) = \text{InvGamma}\left(\frac{n}{2}, \frac{n\hat{\sigma}^2}{2}\right)$ $\leftarrow$ Requires θ to compute

  $$\left\{ \begin{array}{l} \text{if } Y \sim \text{InvGamma}(\alpha, \beta), \\ \text{Var}[Y] = \dfrac{\beta^2}{(\alpha-1)^2(\alpha-2)} \end{array} \right\}$$

Note :

$$\hat{\sigma}^2_{MLE} = \frac{\Sigma(X_i-\theta)^2}{n} = \frac{SSE}{n}$$

$$n\hat{\sigma}^2 = SSE$$

- 2nd/ $\underbrace{P(\theta,\sigma^2|X)}_{\text{NormInvGam}(\dots)} = \underbrace{P(\sigma^2|\theta,X)}_{\text{InvGamma}\left(\frac{n}{2},\frac{n\hat{\sigma}^2}{2}\right)} \; P(\theta|X)$

Marginal $\theta|X$

- $P(\theta|X) = \int_0^\infty P(\theta,\sigma^2|X)\,d\sigma^2$ (uncertainty of σ^2 is removed)

- $$P(\theta|X) = \frac{P(\theta,\sigma^2|X)}{P(\sigma^2|\theta,X)} \propto \frac{\left(\sigma^2\right)^{-n/2} e^{-\frac{1}{2\sigma^2}\overbrace{\Sigma(X_i-\theta)^2}^{=n\hat{\sigma}^2}} \overbrace{(\sigma^2)^{-1}}^{\text{Jeffrey}}}{\frac{\left(\frac{n\hat{\sigma}^2}{2}\right)^{\frac{n}{2}}}{\Gamma(n/2)} \sigma^{2^{-\left(\frac{n}{2}\right)-1}} e^{-\left(\frac{n\hat{\sigma}^2}{2}\right)\frac{1}{\sigma^2}}} \propto \frac{e^{-\frac{n\hat{\sigma}^2}{2\sigma^2}}}{\left(\frac{n\hat{\sigma}^2}{2}\right)^{\frac{n}{2}} e^{-\frac{n\hat{\sigma}^2}{2\sigma^2}}}$$

$$P(\theta|X) = \left(\frac{n\hat{\sigma}^2}{2}\right)^{-n/2} = \left(\frac{SSE}{2}\right)^{-n/2} = \left(\frac{\Sigma(X_i-\theta)^2}{2}\right)^{-n/2} = \left(\frac{(n-1)s^2}{2} + 0 + \frac{n(\bar{X}-\theta)^2}{2}\right)^{-n/2} \times \underbrace{\left(\frac{1}{\left(\frac{(n-1)s^2}{2}\right)}\right)^{-n/2}}_{\text{constant}}$$

- $$\boxed{P(\theta|X)} = \left(1 + \frac{\frac{n(\bar{X}-\theta)^2}{\cancel{2}}}{\frac{(n-1)s^2}{\cancel{2}}}\right)^{-\frac{n}{2}} = \left(1 + \frac{1}{(n-1)}\left(\frac{\bar{X}-\theta}{\frac{s}{\sqrt{n}}}\right)^2\right)^{-\frac{n}{2}} \propto \boxed{T_{n-1}\left(\bar{X}, \frac{s}{\sqrt{n}}\right)}$$

- Sampling similarly:
  1) θ_0 from $T_{n-1}\left(\bar{X}, \frac{s}{\sqrt{n}}\right)$
  2) σ_0^2 from $\text{InvGamma}\left(\frac{n}{2}, \frac{\Sigma(X_i-\theta_0)^2}{2}\right)$
  3) return values $\langle \theta_0, \sigma_0^2 \rangle$

* Note $X_1 \dots X_n | \theta, \sigma^2 \sim N(\theta,\sigma^2)$

$$\frac{\bar{X}-\theta}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

- Let $V \sim T_n := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{V^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}$ } Standard T distribution

- if $W = \sigma V + \mu = f(v)$; $v = f^{-1}(W) = \frac{W-\mu}{\sigma}$

if σ is unknown

$$f_W(w) = f_V\left(f^{-1}(w)\right)\left|\frac{d}{dw}\left[f^{-1}(w)\right]\right|$$

$$\frac{\bar{X}-\theta}{\frac{s}{\sqrt{n}}} \sim \;?$$

$$= \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{\left(\frac{w-\mu}{\sigma}\right)^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} \cdot \frac{1}{\sigma}$$

- $? = T_n(\mu,\sigma^2) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n \sigma^2}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{1}{n}\left(\frac{w-\mu}{\sigma}\right)^2\right)^{-\left(\frac{n+1}{2}\right)}$ } non-standard T dist.