Version 1

Facts:

Beta function $B(x,y) = \frac{(x-1)!\,(y-1)!}{(x+y-1)!}$

Beta dist $\pi(\theta; x, y) = \frac{1}{B(x,y)} \theta^{x-1}(1-\theta)^{y-1}$

useful Integral: $\int_{\theta=0}^{1} \theta^{r}(1-\theta)^{n-r}\, d\theta = \frac{r!\,(n-r)!}{(n+1)!}$

Slot machine has bernoulli model for one try and hence binomial for n trials. ~~If r are successes, then~~

Let r be # of successes.

Let's use Beta dist $\pi(\theta; \alpha, \beta)$ as prior.

$f_{\theta}(\theta) = \pi(\theta; \alpha, \beta)$

$P(X|\theta) = {}^{n}C_{r}\, \theta^{r}(1-\theta)^{n-r}$

To understand this consider θ taking a discrete set of values with some probability distribution

②

JPT for Bernoulli

θ	$P_\theta(\theta)$
0.2	0.1
0.5	0.2
0.7	0.7
0.9	0.0

* $x = 1$ win
  $= 0$ lose

θ	$p(\theta)$
0.2	(0.1)

$x \backslash \theta$	0.2	0.5	0.7	0.9
0	1-0.2	1-0.5	1-0.7	1-0.9
1				

$x \backslash \theta$	0.2	0.5	0.7	0.9
0	1-0.1	1-0.2	1-0.7	
1	0.1	0.2	0.7	

$x \backslash \theta$	0.2	0.5	0.7	0.9
0	(1-0.2)(0.1)	(1-0.5)(0.2)	(1-0.7)(0.7)	(1-0.9)(0)
1	(0.2)(0.1)	(0.5)(0.2)	(0.7)(0.7)	0.9×0 = 0
Total	0.1	0.2	0.7	

so for ~~one~~ a attempt, success is (3pg) | failure is

$P(r=0|\theta) = (1-\theta)\, P(\theta)$ $\quad P_{r|\theta}$

$\cancel{P(\theta}\ \ P_{r|\theta}(r=1|\theta) = \theta\, f_\theta(\theta)$

For r ~~attempts~~

$$f_{R|\theta}(r|\theta) = \theta^r (1-\theta)^{n-r}\, \cancel{p_\theta(\theta)}$$

So, likelihood for r wins

$$f_{r|\theta}(r|\theta) = {}^nC_r\, \theta^r (1-\theta)^{n-r}\, \cancel{f_\theta(\theta)}$$

$$f_R(r) = \int_\theta f_{r|\theta}(r|\theta)\, f_\theta(\theta)\, d\theta$$

$$= \int_\theta {}^nC_r\, \theta^r (1-\theta)^{n-r} \frac{1}{B(\alpha,\beta)}\, \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

$$= {}^nC_r \frac{1}{B(\alpha,\beta)} \int_{\theta=0}^{1} \theta^{r+\alpha-1} (1-\theta)^{n-r+\beta-1}\, d\theta$$

$$= {}^nC_r \frac{1}{B(\alpha,\beta)} \frac{(r+\alpha-1)!\,(n-r+\beta-1)!}{(n+\alpha+\beta-1)!}$$

(4)

$$f_{\theta|R}(\theta|r) = \frac{f_{R|\theta}(r|\theta)\, f_\theta(\theta)}{f_R(r)}$$

$$= \frac{{}^nC_r\, \theta^r (1-\theta)^{n-r} \frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}}{{}^nC_r \frac{1}{B(\alpha,\beta)} \frac{(r+\alpha-1)!\,(n-r+\beta-1)!}{(n+\alpha+\beta-1)!}}$$

$$= \theta^{r+\alpha-1} (1-\theta)^{n-r+\beta-1} \frac{(n+\alpha+\beta-1)!}{(r+\alpha-1)!\,(n-r+\beta-1)!}$$

$$= \frac{\theta^{x-1} (1-\theta)^{y-1}}{\frac{(x-1)!\,(y-1)!}{(x+y-1)!}}$$

where $x = r+\alpha$
$y = n-r+\beta$

$$= \pi(r+\alpha,\ n-r+\beta)$$

Pg 5

$$p(\overset{r_{n+1}}{\check{x}} \mid R) = \int_\theta f_{R,\theta}(\tilde{x}, \theta \mid x)\, d\theta$$

$$= \int f_{R|\theta}(\tilde{x} \mid \theta, R)\, f_{\theta|R}(\theta \mid R)\, d\theta$$

$$= \int \cancel{f_{R|\theta}}\, f_{R|\theta}(\check{x} \mid \theta)\, f_{\theta|R}(\theta \mid R)\, d\theta$$

① Markov assumption

$$= \int \theta \; \pi(r+\alpha, n-r+\beta)\, d\theta$$

$$= E[\theta \mid R] = \frac{r+\alpha}{r+\alpha+n-r+\beta}$$

$$= \frac{r+\alpha}{n+\alpha+\beta}$$

For uniform prior $\alpha = \beta = 1$

$$p(\check{x} \mid R) = \frac{r+1}{n+2} = \frac{1+1}{5+2} = \frac{2}{7}$$

Eqn (1) can be used other ways too:

For ML estimate

$$f_{\tilde{r}|R}(\tilde{r}|R) = \int_\theta f_{R|\theta}(\tilde{r}|\theta)\, f_{\theta|R}(\theta|r)\, d\theta$$

$$\approx \int_\theta f_{R|\theta}(\tilde{r}|\hat{\theta}_{ML})\, f_{\theta|R}(\theta|r)\, d\theta$$

$$= f_{\tilde{r}|\theta}(\tilde{r}|\hat{\theta}_{ML}) \underbrace{\int_\theta f_{\theta|R}(\theta|r)\, d\theta}_{=1}$$

Or for MAP, the same

$$f_{\tilde{r}|R}(\tilde{r}|R) \approx f_{R|\theta}(\tilde{r}|\hat{\theta}_{MAP})$$

Note: In Eqn (1) we said markov assumption (which is essentially markov property). But since all data is conditionally independent given θ

$P(x|\theta, D_n) = P(x|\theta)$