UC Berkeley
Department of Electrical Engineering and Computer Sciences

EE126: Probability and Random Processes

Problem Set 6 Solutions
Spring 2011

Solution to Problem 1. ...

(a) We should have

$$1 = \int_{-\infty}^{\infty} f_R(r)\, dr = \gamma \int_1^3 r dr = \gamma \frac{r^2}{2}\Big|_1^3 = \gamma \frac{3^2 - 1}{2},$$

implying $\gamma = \frac{1}{4}$.

(b) The CDF is defined as $F_R(r) := \mathbb{P}(R \leq r)$. Clearly, $F_R(r) = 0$ for $r \leq 1$, and $F_R(r) = 1$ for $r \geq 3$. For $1 < r < 3$, we have

$$F_R(r) = \int_{-\infty}^{r} f_R(s) ds = \int_1^r \frac{s}{4}\, ds = \frac{s^2}{8}\Big|_1^r = \frac{r^2 - 1}{8}.$$

A compact way of writing $F_R(\cdot)$ is

$$F_R(r) = \max\Big\{ \min\Big\{\frac{r^2 - 1}{8}, 1\Big\}, 0\Big\}, \quad r \in \mathbb{R}.$$

(c) We have

$$\mathbb{E}(R) = \int_{-\infty}^{\infty} r f_R(r)\, dr = \int_1^3 \frac{r^2}{4}\, dr = \frac{r^3}{12}\Big|_1^3 = \frac{13}{6},$$

and

$$\mathbb{P}(A) = \int_A f_R(r) dr = \int_2^3 \frac{r}{4}\, dr = \frac{r^2}{8}\Big|_2^3 = \frac{5}{8}.$$

The conditional density of R given $A = \{R \geq 2\}$ is

$$f_{R|A}(r|A) = \begin{cases} \frac{f_R(r)}{\mathbb{P}(A)} & r \in A \\ 0, & \text{otherwise} \end{cases} = \begin{cases} \frac{2r}{5} & r \in [2, 3] \\ 0, & \text{otherwise} \end{cases},$$

hence,

$$\mathbb{E}(R \mid A) = \int_{-\infty}^{\infty} r f_{R|A}(r|A)\, dr = \int_2^3 \frac{2r^2}{5}\, dr = \frac{2r^3}{15}\Big|_2^3 = \frac{38}{15}.$$

(d) We have

$$\mathbb{E}(W) = \int_{-\infty}^{\infty} r^2 f_R(r)\,dr = \int_1^3 \frac{r^3}{4}\,dr = \frac{r^4}{16}\Big|_1^3 = 5.$$

To compute the variance, we first compute the second moment

$$\mathbb{E}(W^2) = \mathbb{E}(R^4) = \int_{\infty}^{\infty} r^4 f_R(r)\,dr = \int_1^3 \frac{r^5}{4} = \frac{r^6}{24}\Big|_1^3 = \frac{91}{3}.$$

Now,

$$\text{var}(W) = \mathbb{E}(W^2) - (\mathbb{E}W)^2 = \frac{91}{3} - 25 = \frac{16}{3}.$$

(e) Let us first obtain the CDF of Z. Note that using the indicator function notation, we can express the PDF of R as $f_R(r) = \frac{1}{4} r 1\{1 < r \leq 3\}$. We have, for $z > 0$,

$$\begin{aligned}\mathbb{P}(Z \leq z) = \mathbb{P}\big(e^{-R} \leq z\big) &= \mathbb{P}(-R \leq \log z) \\ &= \mathbb{P}(R \geq -\log z) \\ &= \int_{-\log z}^{\infty} \frac{r}{4} 1\{1 < r \leq 3\}\,dr.\end{aligned}$$

Note that if $-\log z \leq 1$, the integral evaluates to 1 and if $-\log z > 3$, the integral evaluates to 0. For $-\log z \in (1, 3]$, we have

$$\mathbb{P}(Z \leq z) = \int_{-\log z}^{3} \frac{r}{4} dr = \frac{9 - (-\log z)^2}{8}.$$

Taking derivative we obtain the PDF

$$f_Z(z) = -\frac{\log z}{4z} 1\{e^{-3} < z \leq e^{-1}\}.$$

Another way of writing this is

$$f_Z(z) = \frac{z^{-1} \log z^{-1}}{4} 1\{e \leq z^{-1} < e^3\}$$

which clearly shows that f_Z is nonnegative.

(f) Let us first determine the PDF of $R := \sqrt{X^2 + Y^2}$ via its CDF. We have, for $r > 0$,

$$\mathbb{P}(R \leq r) = \mathbb{P}(X^2 + Y^2 \leq r^2) = \begin{cases} 0 & r^2 \leq a_0^2 \\ \frac{\pi(r^2 - a_0^2)}{\pi(a_1^2 - a_0^2)} & a_0^2 < r^2 \leq a_1^2 \\ 1 & r^2 \leq a_1^2. \end{cases}$$

Taking derivative we obtain

$$f_R(r) = \frac{2}{a_1^2 - a_0^2} r 1\{a_0 < r \leq a_1\}.$$

Hence it is indeed possible to obtain the desired PDF by setting $a_0 = 1$ and $a_1 = 3$. The claim is true.

Solution to Problem 2. ...

(a) Let T_a be the time it takes for Alice to do the problem set. We know that $\mathbb{P}(T_a > t) = \exp(-\lambda_a t)$, $t > 0$ where $\lambda_a = \frac{1}{\mathbb{E}(T_a)} = \frac{1}{4}$. We get

$$\mathbb{P}(T_a > 4) = \exp(-4\lambda_a) = e^{-1}.$$

(b) We could compute the conditional probability as usual. However, recalling the memoryless property of the exponential distribution, we see that the conditional probability that the problem set will take more than 8 hours given that it has already taken 4, is just the answer to part (a).

(c) Denoting the time it takes for Bob to finish as T_b and noting that T_b is exponential with rate $\lambda_b = \frac{1}{\mathbb{E}(T_b)} = \frac{1}{6}$, we have

$$\begin{aligned}\mathbb{P}(T_a < T_b) = \int_{0<t<s} f_{T_a}(t) f_{T_b}(s)\, dt\, ds &= \int_0^\infty \lambda_a e^{-\lambda_a t} \int_t^\infty \lambda_b e^{-\lambda_b s} ds\, dt \\ &= \int_0^\infty \lambda_a e^{-\lambda_a t} e^{-\lambda_b t} dt = \frac{\lambda_a}{\lambda_a + \lambda_b} = \frac{3}{5}.\end{aligned}$$

(d) This event is the same as $\{T_b > T_a + 1\} \cup \{T_a > T_b + 1\} = \{|T_a - T_b| > 1\}$. Let us consider

$$\begin{aligned}\mathbb{P}(T_b > T_a + 1) = \int_{0<t<s-1} f_{T_a}(t) f_{T_b}(s)\, dt\, ds &= \int_0^\infty \lambda_a e^{-\lambda_a t} \int_{t+1}^\infty \lambda_b e^{-\lambda_b s} ds\, dt \\ &= \int_0^\infty \lambda_a e^{-\lambda_a t} e^{-\lambda_b (t+1)} dt = \frac{\lambda_a}{\lambda_a + \lambda_b} e^{-\lambda_b}.\end{aligned}$$

By symmetry, the result for the other event is similar; one just changes a subscripts to b and vice versa. Hence

$$\mathbb{P}(|T_a - T_b| > 1) = \frac{\lambda_a}{\lambda_a + \lambda_b} e^{-\lambda_b} + \frac{\lambda_b}{\lambda_a + \lambda_b} e^{-\lambda_a} = \frac{3}{5} e^{-\frac{1}{6}} + \frac{2}{5} e^{-\frac{1}{4}}.$$

Solution to Problem 3. ...

(a) The region $\{(x, y) :\ 0 \le y \le x \le 2\}$ is a triangle of area 2 in xy-plane. We have

$$f_{X,Y}(x, y) = \frac{1}{2} 1\{0 \le y \le x \le 2\}.$$

We have

$$\begin{aligned}\mathbb{E}(X) &= \int\int x f_{X,Y}(x, y) dx dy = \int_0^2 \frac{x}{2} \int_0^x dy\, dx = \int_0^2 \frac{x^2}{2} dx = \frac{x^3}{6}\Big|_0^2 = \frac{4}{3}, \\ \mathbb{E}(Y) &= \int\int y f_{X,Y}(x, y) dx dy = \int_0^2 \int_0^x \frac{y}{2} dy\, dx = \int_0^2 \frac{x^2}{4} dx = \frac{2}{3},\end{aligned}$$

and similarly,

$$\mathbb{E}(T) = \int\int (x + y) f_{X,Y}(x, y) dx dy = \int_0^2 \int_0^x \frac{x + y}{2} dy\, dx = \int_0^2 \frac{3}{4} x^2\, dx = \frac{x^3}{4}\Big|_0^2 = 2.$$

(b) $\sigma_T^2 > \sigma_X^2 + \sigma_T^2$ since X and Y have a positive covariance (i.e., X and Y are positively correlated).

Solution to Problem 4. Recalling problem (3)a of HW5, we write, for $t > 0$

$$\mathbb{P}\big(\min\{X_1, \dots, X_n\} > t\big) = \prod_{i=1}^n \mathbb{P}(X_i > t) = \prod_{i=1}^n e^{-\lambda_i t} = e^{-(\sum_{i=1}^n \lambda_i)t},$$

Hence the distribution of Y is exponential with rate $\lambda := \sum_{i=1}^n \lambda_i$. The PDF is

$$f_Y(y) = \lambda e^{-\lambda y} 1\{y \geq 0\}.$$

Solution to Problem 5. ...

(a) Let $\gamma(z)$ denote the PDF of a standard normal random variable, that is,

$$\gamma(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$

Let G and B be the events of good and bad weather, respectively. Conditioned on G, Z is $N(s, 1)$, that is, $f_Z(z|G) = \gamma(z - s)$. Similarly, $f_Z(z|B) = \frac{1}{2}\gamma(\frac{z-s}{2})$. Hence, using $s = 2$,

$$f_Z(z) = \frac{1}{2}\big\{f_Z(z|G) + f_Z(z|B)\big\} = \frac{1}{2}\big\{\gamma(z-2) + \frac{1}{2}\gamma(\frac{z-2}{2})\big\}.$$

(b) Let $\Phi(t) := \int_{-\infty}^t \gamma(z)dz$ be the CDF of standard normal distribution. Note that for $r > 0$, we have

$$\int_{-r}^r \gamma(z)dz = 2\int_0^r \gamma(z)dz = 2\Big\{\int_{-\infty}^r \gamma(z)dz - \frac{1}{2}\Big\} = 2\Phi(r) - 1.$$

The desired probability is

$$\begin{aligned}
\mathbb{P}\big(Z \in (1,3)\big) = \int_1^3 f_Z(z)dz &= \frac{1}{2}\Big\{\int_1^3 \gamma(z-2)\,dz + \int_1^3 \frac{1}{2}\gamma(\frac{z-2}{2})\,dz\Big\} \\
&= \frac{1}{2}\Big\{\int_{-1}^1 \gamma(z)\,dz + \int_{-\frac{1}{2}}^{\frac{1}{2}} \gamma(z)\,dz\Big\} \\
&= \frac{1}{2}\Big\{2\Phi(1) - 1 + 2\Phi(\tfrac{1}{2}) - 1\Big\} = \Phi(1) + \Phi(\tfrac{1}{2}) - 1 \approx 0.5328.
\end{aligned}$$

Solution to Problem 6. ...
We can write, breaking the integral over pieces where $\lfloor x \rfloor$ is constant,

$$\begin{aligned}
\mathbb{E}(\lfloor X \rfloor) = \int \lfloor x \rfloor f_X(x)dx &= \int_0^\infty \lfloor x \rfloor e^{-x}\,dx \\
&= \sum_{n=0}^\infty \int_n^{n+1} ne^{-x}dx = \sum_{n=0}^\infty n(e^{-n} - e^{-(n+1)}) = (1 - e^{-1})\underbrace{\sum_{n=0}^\infty ne^{-n}}_{=:S}.
\end{aligned}$$

The infinite sum can be evaluated in different ways. For example,

$$S = \sum_{n=0}^{\infty}\sum_{k=1}^{n} e^{-n} = \sum_{n=0}^{\infty}\sum_{k=1}^{\infty} e^{-n}1\{k \le n\} = \sum_{k=1}^{\infty}\sum_{n=0}^{\infty} e^{-n}1\{k \le n\}$$
$$= \sum_{k=1}^{\infty}\sum_{n=k}^{\infty} e^{-n} = \sum_{k=1}^{\infty} \frac{e^{-k}}{1-e^{-1}} = \frac{e^{-1}}{(1-e^{-1})^2}.$$

As another example, we might look at $f(x) = \sum_{n=0}^{\infty} x^n e^{-n} = \frac{1}{1-xe^{-1}}$, for $|xe^{-1}| < 1$. Taking derivatives, once term-by-term, and once of the closed form, we get

$$\frac{e^{-1}}{(1-xe^{-1})^2} = f'(x) = \sum_{n=1}^{\infty} n x^{n-1} e^{-n}.$$

Letting $x = 1$ in the above, i.e., looking at $f'(1)$ gives the desired result.
In any case, we obtain

$$\mathbb{E}(\lfloor X \rfloor) = \frac{e^{-1}}{1-e^{-1}}.$$

Solution to Problem 7. ...

(a) First note that by letting $Z = X/\sigma$, we can work with a standard normal random variable. That is, we compute $\mathbb{E}Z^n$; then, $\mathbb{E}X^n = \sigma^n \mathbb{E}Z^n$. Recalling the density of a standard normal, we observe that by symmetry around zero, the odd moments are zero. [We are using the fact that $\int_{-\infty}^{\infty} g(x)\,dx = 0$, when g is an odd function such that $\int_0^\infty |g(x)| dx < \infty$.]

For even moments, i.e., even n, we write

$$\begin{aligned} m_n := \mathbb{E}(Z^n) = 2\int_0^\infty z^n \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz &= \frac{2}{\sqrt{2\pi}} \int_0^\infty z^{n-1} z e^{-z^2/2} dz \\ &= \frac{2}{\sqrt{2\pi}} \int_0^\infty z^{n-1} d(-e^{-z^2/2})\,dz \\ &= \frac{2}{\sqrt{2\pi}} \Big\{ -z^{n-1} e^{-z^2/2}\Big|_0^\infty + \int_0^\infty (n-1) z^{n-2} e^{-z^2/2}\,dz \Big\} \\ &= (n-1)\frac{2}{\sqrt{2\pi}} \int_0^\infty z^{n-2} e^{-z^2/2}\,dz \\ &= (n-1) m_{n-2}. \qquad (1) \end{aligned}$$

Noting that $m_2 = \mathbb{E}Z^2 = \mathrm{var}(Z) = 1$, by assumption, we can use the recursion above to obtain $m_n = (n-1)(n-3)\cdots 3 \cdot 1$ for even n. That is, m_n is the product of odd integers from 1 to $n-1$. This is usually denoted compactly as $(n-1)!!$. Thus, we have

$$\mathbb{E}(X^n) = \begin{cases} \sigma^n (n-1)!! & n \text{ even} \\ 0 & n \text{ odd} \end{cases}.$$

It is also worth noting that the above recursion, i.e. (1), holds for absolute odd moments $\mathbb{E}|Z|^n$, n odd. [Check this for yourself. You do not even need to modify the argument.] In particular, if we compute $\mathbb{E}|Z|$, we get (another) initial point for the recursion and we get all the odd absolute moments. As we will need $\mathbb{E}|Z|$ for a later part, let us do it here. We have

$$\mathbb{E}|Z| = 2\int_0^\infty z\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz = \frac{2}{\sqrt{2\pi}}\Big[-e^{-z^2/2}\Big|_0^\infty = \sqrt{\frac{2}{\pi}}. \tag{2}$$

(b) By linearity of the expectation operator, we have

$$\mathbb{E}(Y) = \sum_{i=1}^n \mathbb{E}(X_i^2) = n\sigma^2.$$

To compute the variance, we use the property that for a sum of "independent" random variables, the variance of the sum is the sum of variances. Hence,

$$\begin{aligned}\mathrm{var}(Y) = \sum_{i=1}^n \mathrm{var}(X_i^2) = n\,\mathrm{var}(X_1^2) &= n\big[\mathbb{E}(X_1^4) - (\mathbb{E}X_1^2)^2\big] \\ &= n\sigma^4\big[3\cdot 1 - 1^2\big] = 2n\sigma^4.\end{aligned}$$

where we have used the result of part (a) to compute the fourth moment of X_1. It is worth mentioning that, when $\sigma = 1$, the random variable Y is said to have (central) chi-square distribution with n degrees of freedom, denoted as χ_n^2.

(c) By linearity of expectation and (2) we have

$$\mathbb{E}(W) = \sum_{i=1}^n \mathbb{E}|X_i| = n\mathbb{E}|X_1| = \sqrt{\frac{2}{\pi}}n\sigma.$$

For the variance we get, as in part (b),

$$\begin{aligned}\mathrm{var}(W) = \sum_{i=1}^n \mathrm{var}(|X_i|) = n\,\mathrm{var}(|X_1|) &= n\big[\mathbb{E}(|X_1|^2) - (\mathbb{E}|X_1|)^2\big] \\ &= n\sigma^2\big[1 - \big(\sqrt{\tfrac{2}{\pi}}\big)^2\big] = n\sigma^2\big(1 - \frac{2}{\pi}\big).\end{aligned}$$

(d) As suggested by the hint, one approach is to use the convexity of $g(x) = -\sqrt{x}$ together with Jensen's inequality. We get $\mathbb{E}g(Y) \geq g(\mathbb{E}Y)$ or after multiplication of both sides by -1,

$$\mathbb{E}\sqrt{Y} \leq \sqrt{\mathbb{E}Y} = \sqrt{n}\sigma.$$

(e) In this part we are assuming $\sigma = 1$. There are different approaches to solving this problem. Here is one. Note that, for $y > 0$,

$$\begin{aligned}\int_0^y f_Y(t)dt &= \mathbb{P}(Y \le y) \\ &= \mathbb{P}(X_1^2 + X_2^2 \le y) \\ &= \int_{x_1^2+x_2^2 \le y} \frac{1}{2\pi} e^{-(x_1^2+x_2^2)/2}\, dx_1 dx_2 \\ &= \int_0^{2\pi} \int_0^{\sqrt{y}} \frac{1}{2\pi} e^{-r^2/2} r dr d\theta \\ &= \int_0^{\sqrt{y}} r e^{-r^2/2} dr.\end{aligned}$$

Differentiating both sides with respect to y, we get for $y > 0$,

$$f_Y(y) = (\sqrt{y})' \left[re^{-r^2/2} \right|_{r=\sqrt{y}} = \frac{1}{2\sqrt{y}} \sqrt{y} e^{-y/2}.$$

Hence, for $n = 2$,

$$f_Y(y) = \frac{1}{2} e^{-y/2} 1\{y > 0\},$$

which is the exponential distribution with mean 2.

This argument may be generalized to all $n > 2$, once you know how to do a polar change of variables in $\mathbb{R}^n$. For any vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, let $\|x\| := \sqrt{x_1^2 + x_2^2 + \cdot + x_n^2}$ denote its Euclidean norm, measuring the distance of x from origin. Consider the sphere $S^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}$, that is all the points that are at distance 1 from the origin. Also consider the unit ball $B_n := \{x \in \mathbb{R}^n : \|x\| \le 1\}$. Let v_n denote the volume of the unit ball B_n.

Let $B_n(r) := \{x \in \mathbb{R}^n : \|x\| \le r\}$. Convince yourself that the volume of $B_n(r)$ should be $v_n r^n$.

Any vector $x \in \mathbb{R}^n$ can be described as $x = r\theta$ where $r = \|x\| \ge 0$ and θ lies on the sphere S^{n-1}, hence $\|\theta\| = 1$.

In order to do integrals in (r, θ) coordinates, we need to figure out the element of volume located at (r, θ). Let $\widetilde{\sigma}_{n-1}(d\theta)$ be the area of a tiny patch $d\theta$ on S^{n-1} located around θ. (Area is not strictly correct as we are dealing with an $(n-1)$-dimensional object but let us use it in analogy with 3-dimensional space.) Then since the sphere is $(n-1)$-dimensional, the area of a patch located at θ on rS^{n-1} (that is, sphere of radius r) is $r^{n-1}\widetilde{\sigma}_{n-1}(d\theta)$. Hence the element of volume located at (r, θ) is $r^{n-1}\widetilde{\sigma}_{n-1}(d\theta)dr$.

It is customary to normalize $\widetilde{\sigma}_{n-1}$ so that the area of the unit sphere is one. Note that the area of the unit sphere is just $\widetilde{\sigma}_{n-1}(S^{n-1}) = \int_{S^{n-1}} \widetilde{\sigma}_{n-1}(d\theta)$. On the other hand, we can get the area of rS^{n-1} (a sphere of radius r) by the usual trick of looking at the volume of a shell located between r and $r + dr$, that is,

$$d(v_n r^n) = v_n n r^{n-1} dr$$

where the factor multiplying dr is the area of rS^{n-1}. Letting $r = 1$, we conclude that the area of S^{n-1}, that is $\widetilde{\sigma}_{n-1}(S^{n-1})$, is nv_n. Letting

$$\sigma_{n-1} := (nv_n)^{-1}\widetilde{\sigma}_{n-1},$$

we get our normalized measure of area on S^{n-1}. That is, $\sigma_{n-1}(S^{n-1}) = 1$.

Now an integral in polar coordinate can be written as

$$\int_{\mathbb{R}^n} f(x)dx = nv_n \int_0^\infty \int_{S^{n-1}} f(r\theta)r^{n-1}\sigma_{n-1}(d\theta)\,dr.$$

Equipped with this, we go back to our problem and write (using $X := (X_1, \ldots, X_n)$ so that $Y = \|X\|^2$)

$$\begin{aligned}
\int_0^y f_Y(t)dt &= \mathbb{P}(Y \le y) \\
&= \mathbb{P}(\|X\|^2 \le y) \\
&= \int_{\|x\| \le \sqrt{y}} \frac{1}{(2\pi)^{n/2}} e^{-\|x\|^2/2} dx \\
&= \frac{nv_n}{(2\pi)^{n/2}} \int_{r \le \sqrt{y}} \int_{S^{n-1}} e^{-r^2/2} r^{n-1} \sigma_{n-1}(d\theta) dr \\
&= \frac{nv_n}{(2\pi)^{n/2}} \int_0^{\sqrt{y}} r^{n-1} e^{-r^2/2} dr \underbrace{\int_{S^{n-1}} \sigma_{n-1}(d\theta)}_{=\sigma_{n-1}(S^{n-1})=1}.
\end{aligned}$$

Differentiating we get, for $y > 0$

$$f_Y(y) = \frac{1}{2\sqrt{y}} \frac{nv_n}{(2\pi)^{n/2}} y^{(n-1)/2} e^{-y/2}.$$

Hence,

$$f_Y(y) = \frac{nv_n}{2(2\pi)^{n/2}} y^{\frac{n}{2}-1} e^{-y/2} 1\{y > 0\}.$$

We are done and the leading factor is just a constant. But since we know that f_Y should integrate to 1, we can express the factor in terms of an integral and get an expression for v_n for free. We have

$$\frac{2(2\pi)^{n/2}}{nv_n} = \int_0^\infty y^{\frac{n}{2}-1} e^{-y/2} dy = 2^{\frac{n}{2}} \int_0^\infty z^{\frac{n}{2}-1} e^{-z} dz = 2^{\frac{n}{2}} \Gamma(n/2),$$

hence

$$v_n = \frac{\pi^{n/2}}{(n/2)\Gamma(n/2)} = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}.$$

Note also that

$$f_Y(y) = \frac{2^{-\frac{n}{2}}}{\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-y/2} 1\{y > 0\}$$

which is just Gamma distribution with parameters $n/2$ and $1/2$.

Solution to Problem 8. ...

(a) Let us consider a general setup suggested by the problem, which will work for both parts (a) and (b). Let $p(x|\theta) = \mathbb{P}(X = x \mid \theta)$ be the conditional PMF of X given that the probability of landing heads up is θ. Formally we should write something like $p_{X|\Theta}(x|\theta)$ but me omit the subscript for brevity. Then,

$$p(x|\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x},$$

for $x = 1, \ldots, n$. We also assume Θ to be distributed as $\mathcal{B}e(\alpha, \beta)$. Then, the PDF of Θ is $f(\theta) \propto_\theta \theta^{\alpha-1}(1-\theta)^{\beta-1}$ for $\theta \in (0,1)$, where $\propto_\theta$ means "proportional in θ"; that is, we allow ourselves to drop factors constant in θ. We can obtain the posterior using the Bayes rule

$$\begin{aligned} f(\theta|x) = \frac{p(x|\theta)f(\theta)}{p(x)} &\propto_\theta p(x|\theta)f(\theta) \\ &\propto_\theta \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}. \end{aligned} \tag{3}$$

This shows that the posterior distribution of Θ given $X = x$ is again a Beta distribution, but with updated parameters, namely $\mathcal{B}e(x+\alpha, n-x+\beta)$. If necessary, we can use this observation to compute the constant in (3), but we do not need to do that. Also, knowing the mean of a Beta distribution, we can compute the posterior mean of Θ as

$$\mathbb{E}(\Theta \mid X = x) = \frac{x+\alpha}{x+\alpha+n-x+\beta} = \frac{x+\alpha}{n+\alpha+\beta}. \tag{4}$$

For part (a), we have $\alpha = \beta = 1$. Hence the posterior distribution is $\mathcal{B}e(x+1, n-x+1)$ and the mean is $\frac{x+1}{n+2}$.
For $x = 3$ and $n = 10$, we get numerical values,

$$\mathcal{B}e(4,8), \quad \frac{4}{12} = \frac{1}{3} \approx 0.3333.$$

For $x = 36$ and $n = 100$, we get

$$\mathcal{B}e(37,65), \quad \frac{37}{102} \approx 0.3627.$$

The first row of Fig. 1 shows the plots of prior/posterior and their means for the two cases. ($n = 10$ is on the left column.)

(b) Using the results of part (a), we get $\mathcal{B}e(x+\frac{1}{2}, n-x+\frac{1}{2})$, $\frac{x+\frac{1}{2}}{n+1}$ for the posterior and its mean. For $x = 3$ and $n = 10$, we get numerical values,

$$\mathcal{B}e(3.5, 7.5), \quad \frac{3.5}{7.5} \approx 0.3182.$$

For $x = 36$ and $n = 100$, we get

$$\mathcal{B}e(36.5, 64.5), \quad \frac{36.5}{64.5} \approx 0.3614.$$

Second row of Fig. 1 shows the plots of this part.

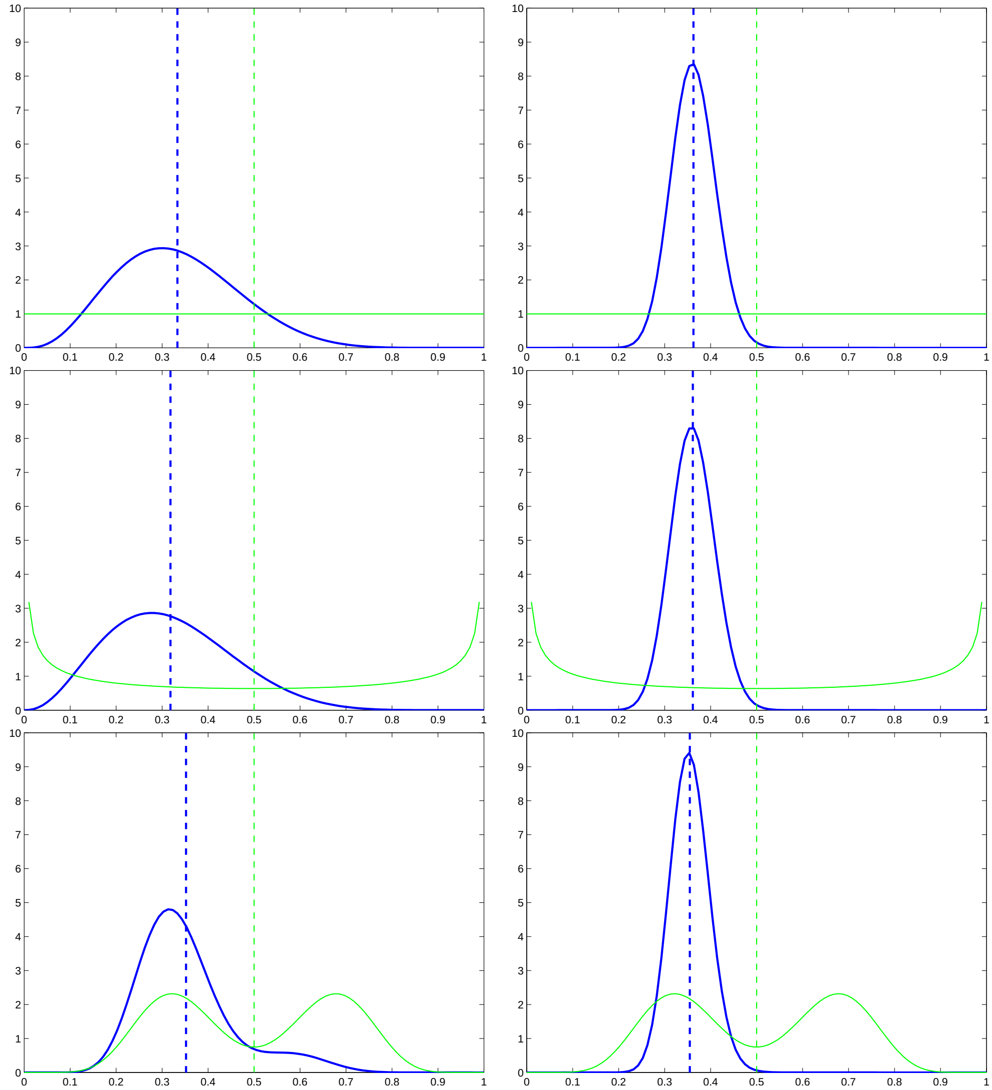

Figure 1: Plots of problem 8

(c) We can express the prior distribution of Θ in this part as a mixture of two Beta distributions, namely

$$\frac{1}{2}\mathcal{B}e(20,10)+\frac{1}{2}\mathcal{B}e(10,20).$$

Let us consider a general mixture of the form

$$\gamma\mathcal{B}e(\alpha_1,\beta_1)+(1-\gamma)\mathcal{B}e(\alpha_2,\beta_2),$$

for the prior. In other words, assume the PDF of the prior to be,

$$f(\theta)=\frac{\gamma}{B(\alpha_1,\beta_1)}\theta^{\alpha_1-1}(1-\theta)^{\beta_1-1}+\frac{1-\gamma}{B(\alpha_2,\beta_2)}\theta^{\alpha_2-1}(1-\theta)^{\beta_2-1},$$

for $\theta \in (0,1)$. As in part (a), we get

$$\begin{aligned} f(\theta|x) &\propto_\theta p(x|\theta)f(\theta) \\ &\propto_\theta \frac{\gamma}{B(\alpha_1,\beta_1)}\theta^{x+\alpha_1-1}(1-\theta)^{n-x+\beta_1-1} + \frac{1-\gamma}{B(\alpha_2,\beta_2)}\theta^{x+\alpha_2-1}(1-\theta)^{n-x+\beta_2-1}. \end{aligned}$$

Thus, the posterior is a again a mixture of Beta distributions with updated parameters. To get the normalizing constant correctly, we integrate the RHS of the above over $(0,1)$ to obtain

$$\gamma \underbrace{\frac{B(x+\alpha_1, n-x+\beta_1)}{B(\alpha_1,\beta_1)}}_{A_1} + (1-\gamma)\underbrace{\frac{B(x+\alpha_2, n-x+\beta_2)}{B(\alpha_2,\beta_2)}}_{A_2}.$$

Dividing by this constant and with a slight bit of algebra, we get that the posterior is the mixture

$$\frac{\gamma A_1}{\gamma A_1 + (1-\gamma)A_2}\mathcal{B}e(x+\alpha_1, n-x+\beta_1) + \frac{(1-\gamma)A_2}{\gamma A_1 + (1-\gamma)A_2}\mathcal{B}e(x+\alpha_2, n-x+\beta_2).$$

Specializing to our case where $\gamma = \frac{1}{2}$, we get that the posterior is

$$\frac{A_1}{A_1+A_2}\mathcal{B}e(x+\alpha_1, n-x+\beta_1) + \frac{A_2}{A_1+A_2}\mathcal{B}e(x+\alpha_2, n-x+\beta_2).$$

with mean

$$\frac{A_1}{A_1+A_2}\frac{x+\alpha_1}{n+\alpha_1+\beta_1} + \frac{A_2}{A_1+A_2}\frac{x+\alpha_2}{n+\alpha_2+\beta_2}.$$

For $x=3$, $n=10$, $\alpha_1=\beta_2=20$ and $\alpha_2=\beta_1=10$ we get

$$0.1085\,\mathcal{B}e(23,17) + 0.8915\,\mathcal{B}e(13,27), \quad 0.3521$$

for posterior and its mean.

For $x=36$, $n=100$, we have

$$0.0119\,\mathcal{B}e(56,74) + 0.9881\,\mathcal{B}e(46,84), \quad 0.3548.$$

Third row of Fig 1 shows the plots of this part. This problem is supposed to illustrate that at small sample sizes (say $n=10$) the choice of the prior affects the inference result, while as the sample size increases (cf. $n=100$) the effect of the prior tends to wash out. Note in part (c) that the data forces the posterior to pick the correct mode of the bimodal prior. (That is, for $n=100$ and $x=36$ for example, the weight of $\mathcal{B}e(46,84)$ which has a mean around 0.35 is emphasized.)