Math 302, PSET 3

(1) (a) Define the function

$$f(x) = \begin{cases} 3x - b & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

Show that there is no value of b for which this is the p.d.f. of some RV X.

(b) Let

$$f(x) = \begin{cases} \frac{1}{2}\cos x & x \in [-b,b] \\ 0 & \text{otherwise} \end{cases}$$

Show that there is exactly one value of b for which this could be the p.d.f. of some RV X.

Solution: a) First, since $f(0) \geq 0$, $b \leq 0$. We also need $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$, so

$$1 = \int_0^1 (3x - b)\,\mathrm{d}x = \frac{3}{2} - b.$$

Thus we have $b = \frac{1}{2}$ which does not satisfy $b \leq 0$, so f is not a density function for any b.

b) We have

$$\int_{-b}^{b} \tfrac{1}{2}\cos x \; dx = \frac{1}{2}(\sin b - \sin(-b)) = \sin b,$$

and this equals 1 if $b = \frac{\pi}{2} + 2\pi k$, where k is any integer. If $k \neq 0$, then the interval $[-b,b]$ would contain points at which $\cos x$ is negative, which is impossible for a p.d.f.. Thus, only $k = 0$ is allowed, and indeed, f is a nonnegative function and has integral 1 with this choice of $b = \pi/2$. It could therefore be the p.d.f. of a random variable.

(2) Let $c > 0$ and $X \sim \mathrm{Unif}[0,c]$. Show that the RV $Y = c - X$ has the same c.d.f. and therefore also the same p.d.f. as X.

Solution: We have

$$\mathbb{P}(Y \leq b) = \mathbb{P}(X \geq c - b) = \begin{cases} 0 & b \leq 0 \\ \int_{c-b}^{c} \frac{1}{c} = \frac{b}{c} & 0 \leq b \leq c \\ 1 & c \leq b \end{cases},$$

which is the same as $\mathbb{P}(X \leq b)$. Since the p.d.f. is the derivative of the cdf, also the p.d.f.'s of X and Y coincide.

(3) Let X be a random variable with p.d.f.

$$f(x) = \begin{cases} cx^{-3} & x > 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find c so that f is a p.d.f.
(b) Compute the c.d.f. of X.
(c) Find $\mathbb{P}(X > 3 | X < 5)$.
(d) Find the median of X, i.e. the value m such that $\mathbb{P}(X > m) = \mathbb{P}(X \leq m)$.
(e) Calculate $\mathbb{E}\sqrt{X}$.

Solution: (a) We must have $\int_{-\infty}^{\infty} f(x)dx = 1$, so $c = 8$. (b)

$$F(b) = \int_{-\infty}^{b} f(x)\; dx = \begin{cases} 0 & b < 2 \\ \int_2^b 8x^{-3} & b \geq 2 \end{cases} = \begin{cases} 0 & b < 2 \\ 1 - 4b^{-2} & b \geq 2 \end{cases}$$

(c)

$$\mathbb{P}(X>3|X<5)=\frac{\mathbb{P}(\{X>3\}\cap\{X<5\})}{\mathbb{P}(X<5)}=\frac{\mathbb{P}(X\in(3,5))}{\mathbb{P}(X<5)}=\frac{F(5)-F(3)}{F(5)}=\frac{64}{189}$$

(d) We need to solve $F(m)=\frac{1}{2}$, which gives $m=2\sqrt{2}$.

(e)

$$\mathbb{E}\sqrt{X}=\int_{-\infty}^{\infty}\sqrt{x}\,f(x)\,dx=\int_{2}^{\infty}\sqrt{x}2x^{-2}\,dx=\frac{4\sqrt{2}}{3}.$$

(4) Let X be an Exp(2) random variable. Find a number a such that $\{X\in[0,1]\}$ is independent of $\{X\in[a,2]\}$.

Solution: If $a<0$ then all probabilities are the same as in the case $a=0$, so we may assume that $0\le a\le 1$ (recall that disjoint events that have positive probability are never independent, and hence we can indeed rule out $a\in(1,2)$). We have

$$\mathbb{P}(X\in[0,1])=F_X(1)-F_X(0)=1-e^{-2},$$

and

$$\mathbb{P}(X\in[a,2])=F_X(2)-F_X(a)=e^{-2a}-e^{-4}.$$

The probability of intersection is

$$\mathbb{P}(X\in[0,1],\,X\in[a,2])=\mathbb{P}(X\in[a,1])]=F_X(1)-F_X(a)=e^{-2a}-e^{-2}.$$

The definition of independence gives the equation

$$e^{-2a}-e^{-2}=(1-e^{-2})(e^{-2a}-e^{-4}),$$

so

$$e^{-2a}=1-e^{-2}(1-e^{-2}),$$

that is,

$$a=-\frac{1}{2}\ln(1-e^{-2}(1-e^{-2}))\approx 0.062.$$

(5) Let X be a standard normal random variable. Compute $\mathbb{E}\,X^n$ for all $n\in\mathbb{N}$.

Solution: Let $I_n=\mathbb{E}(X^n)$, we know that $I_0=1$ and $I_1=0$. Now let $n\ge 2$, we prove a recursion for I_n. Let $\varphi(x)=\frac{1}{2\pi}e^{-x^2/2}$ be the PDF of X. As $\varphi'(x)=-x\varphi(x)$, we can use integration by parts with $u=x^{n-1}$ and $dv=x\varphi(x)$:

$$\begin{aligned}
I_n&=\int_{-\infty}^{\infty}x^n\varphi(x)\,\mathrm{d}x\\
&=\int_{-\infty}^{\infty}x^{n-1}(x\varphi(x))\,\mathrm{d}x\\
&=x^{n-1}(-\varphi(x))|_{-\infty}^{\infty}-\int_{-\infty}^{\infty}(n-1)x^{n-2}(-\varphi(x))\,\mathrm{d}x\\
&=0+(n-1)\int_{-\infty}^{\infty}x^{n-2}\varphi(x)\,\mathrm{d}x\\
&=(n-1)I_{n-2}.
\end{aligned}$$

Let $k!!$ (k semifactorial) denote the product of positive integers from 1 to k which has the same parity as k, so $k!!=k(k-2)(k-4)\dots$ The above recursion implies that $I_n=(n-1)!!I_1=0$ if n is odd, and $I_n=(n-1)!!I_0=(n-1)!!$ if n is even.

If n is odd, then $x^n\varphi(x)$ is odd and integrable on $(-\infty,\infty)$, which proves directly that $\mathbb{E}(Z^n) = \int_{-\infty}^{\infty} x^n \varphi(x)\,\mathrm{d}x = 0$.

(6) You have two dice, one with three sides labeled $0, 1, 2$ and one with 4 sides, labeled $0, 1, 2, 3$. Let X_1 be the outcome of rolling the first die, and X_2 the outcome of rolling the second. The rolls are independent.
(a) What is the joint p.m.f. of (X_1, X_2)?
(b) Let $Y_1 = X_1 \cdot X_2$ and $Y_2 = \max\{X_1, X_2\}$. Make a table for the joint p.m.f. of (Y_1, Y_2).
(c) Are Y_1, Y_2 independent? Compute $\mathrm{Cov}(Y_1, Y_2)$.

Solution: (a) By independence we have $p(x, y) = (1/3)(1/4) = 1/12$ for all $x \in \{0, 1, 2\}$ and $y \in \{0, 1, 2, 3\}$.
(b)

Table 1. The p.m.f. of (Y_1, Y_2) with the marginals.

$Y_1 \downarrow\ Y_2 \rightarrow$	0	1	2	3	p_{Y_1}
0	1/12	1/6	1/6	1/12	1/2
1	0	1/12	0	0	1/12
2	0	0	1/6	0	1/6
3	0	0	0	1/12	1/12
4	0	0	1/12	0	1/12
6	0	0	0	1/12	1/12
p_{Y_2}	1/12	1/4	5/12	1/4	

(c) For the marginal distributions see the margins of the above table. Since

$$\mathbb{P}(Y_1 = 1, Y_2 = 0) = 0 \neq \mathbb{P}(Y_1 = 1)\mathbb{P}(Y_2 = 0),$$

the variables Y_1 and Y_2 are not independent.

(7) A fair die is rolled three times with outcomes X_1, X_2, X_3. Let Y_3 be the maximum of the values obtained.
(a) Show that

$$\mathbb{P}(Y_3 \leq j) = \mathbb{P}(X_1 \leq j)^3$$

for any $j = 1, 2, \ldots, 6$. Use this to find the distribution of Y_3.

(b) Suppose instead we sample n independent random variables $U_1, U_2, \ldots, U_n$ with $\mathrm{Unif}(0, 1)$ distribution, and let M_n be their maximum. Find the PDF of M_n.

(c) Note the typo in the limit – it should have said $1 - e^{-x}$. Show that, for any $x \in \mathbb{R}$, $\mathbb{P}(n \cdot (1 - M_n) \leq x) \rightarrow 1 - e^{-x}$ as $n \rightarrow \infty$.

Solution: (a) Note that $Y_3 \leq j$ if and only if $X_1 \leq 3$ and $X_2 \leq j$ and $X_3 \leq j$. Since X_1, X_2, X_3 are independent with the same CDF,

$$\mathbb{P}(Y_3 \leq j) = \mathbb{P}(X_1 \leq 3, X_2 \leq 3, X_3 \leq 3) = \mathbb{P}(X_1 \leq 3)\mathbb{P}(X_2 \leq 3)\mathbb{P}(X_3 \leq 3) = \mathbb{P}(X_1 \leq 3)^3.$$

(b) By the same reasoning, $\mathbb{P}(M_n \leq x) = \mathbb{P}(U_1 \leq x)^n = x^n$. So the PDf of M_n is $f(x) = nx^{n-1}, x \in [0,1]$.

(c) Just plug in: the given probability can be re-written as $\mathbb{P}(M_n > 1 - x/n) = 1 - (1 - x/n)^n \to 1 - e^{-x}, n \to \infty$.

(8) Compute the moment generating functions of $X \sim \mathrm{Geom}(p)$, $Y \sim \mathrm{Exp}(\lambda)$ and of $Z \sim \mathrm{Poisson}(\mu)$.

Solution:
Case Geometric: Using the definition of the m.g.f. and the geometric series, we get

$$\begin{aligned}\mathbb{E}e^{t\cdot\mathrm{Geom}(p)} &= \sum_{k\geq 1} e^{tk}\cdot\mathbb{P}(\mathrm{Geom}(p)=k)\\ &\sum_{k\geq 1} e^{tk}\cdot p\cdot(1-p)^{k-1}\\ &= p\cdot e^t\cdot\sum_{k\geq 1}\left(e^t\right)^{k-1}\cdot(1-p)^{k-1} = \frac{p\cdot e^t}{1-(1-p)e^t}\end{aligned}$$

Note: We could now compute the mean / variance from this function by taking derivatives at 0. Compare this to the tricks we needed in the lecture to compute the mean of Geometric!

Case Exponential: Using the definition, we have

$$\begin{aligned}\mathbb{E}e^{t\cdot\mathrm{Exp}(\lambda)} &= \int_{-\infty}^{\infty} e^{t\cdot x} f(x)\,\mathrm{d}x\\ &= \int_0^{\infty} e^{t\cdot x}\cdot\lambda e^{-\lambda x}\,\mathrm{d}x\\ &= \begin{cases}\frac{\lambda}{\lambda-t} & t<\lambda\\ \infty & \text{else}\end{cases}\end{aligned}$$

Case Poisson:

$$\begin{aligned}\mathbb{E}e^{t\cdot\mathrm{Pois}(\mu)} &= \sum_{k\geq 1} e^{tk}\cdot\mathbb{P}(\mathrm{Pois}(\mu)=k)\\ &\sum_{k\geq 0} e^{tk}\cdot e^{-\mu}\cdot\mu^k/k! = e^{-\mu}\cdot\sum_{k\geq 0}(e^t\mu)^k/k!\\ &= e^{-\mu}\cdot e^{e^t\mu} = \exp[\mu(e^t-1)].\end{aligned}$$

(9) Proof of the 'law of the unconscious statistician'

(a) Let X be a continuous random variable with p.d.f. $f(x)$ and $g : \mathbb{R} \to \mathbb{R}$ be a strictly increasing function for which the set $A := \{x : g'(x) = 0\}$ is finite. Show that the following is a p.d.f. of $g(X)$:

$$f_{g(X)}(y) = \begin{cases}\frac{f(g^{-1}(y))}{|g'(g^{-1}(y))|} & \text{there exists some } x \in \mathbb{R}\setminus A \text{ s.t. } g(x) = y\\ 0 & \text{otherwise}\end{cases}.$$

(Note that the set $g(A) := \{y \in \mathbb{R} : \text{ s.t. } g(x) = y \text{ for some } x \in A\}$ is finite. The values of a function on a finite set does not affect its integral on any interval. Thus one need not worry about the value of $f_{g(X)}$ on A.)

(b)Let X be a continuous random variable with density function f_X. Let g be a differentiable, strictly increasing or strictly decreasing function for which $\mathbb{E}[g(X)]$ and $|A|$

are finite (where A is as above)). Prove that

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx.$$

Hint: Use part (a). When solving part (b) use the definition for the expectation for the continuous random variable $Y := g(X)$, which is $\mathbb{E}[Y] := \int_{-\infty}^{\infty} y f_Y(y) dy$, where f_Y is the p.d.f. of Y. (Note that we have been using the formula $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$ without a proof. Part (a) allows us to prove this formula in the special case considered at part (b).)

Solution:

(a) First consider b such that $g(x) = b$ for some x. We compute the c.d.f., using that g is strictly increasing (the case that g is strictly decreasing is analogous and is left for the reader).

$$F_{g(X)}(b) = \mathbb{P}(g(X) \leq b) = \mathbb{P}(X \leq g^{-1}(b)) = F_X(g^{-1}(b)).$$

Here, $g^{-1}(b)$ is the inverse function of g (e.g. $g^{-1}(b) = \sqrt{b}$ if $g(x) = x^2$, or $g^{-1}(b) = \arctan b$ if $g(x) = \tan x$). Using the chain rule

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x} F_{g(X)}(x) &= F_X'(g^{-1}(x)) \cdot \frac{\mathrm{d}}{\mathrm{d}x} g^{-1}(x) \\ &= f_X(g^{-1}(x)) \cdot \frac{1}{g'(g^{-1}(x))}, \end{aligned}$$

where we used a theorem about the derivative of the inverse function (since $x = g(g^{-1}(x))$ taking derivative and using the chain rule yields $1 = g'(g^{-1}(x)) \cdot \frac{\mathrm{d}}{\mathrm{d}x} g^{-1}(x)$ and so

$$\frac{\mathrm{d}}{\mathrm{d}x} g^{-1}(x) = \frac{1}{g'(g^{-1}(x))}).$$

Now consider $b \notin g(\mathbb{R}) = \{g(x) : x \in \mathbb{R}\}$. Since g is strictly increasing and continuous (since it is differentiable) we must have that $g(\mathbb{R}) = (c, d)$ for some $-\infty \leq c < d \leq \infty$. For all $b > d$ we have $F_{g(X)}(b) = \mathbb{P}(g(X) \leq b) = 1$ and hence $f_{g(X)}(b) = F_{g(X)}'(b) = 0$ (if a function is constant in a neighborhood of a point, the derivative at that point is zero). Likewise for all $b < c$ we have $F_{g(X)}(b) = \mathbb{P}(g(X) \leq b) = 0$ and hence $f_{g(X)}(b) = F_{g(X)}'(b) = 0$.

(b) Let's consider the case that g is increasing. The other case is analogous (with two extra minus signs that will eventually cancel out). By definition $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} y f_{g(X)}(y) \mathrm{d}y$. Using part (a)

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} y \cdot \frac{f(g^{-1}(y))}{g'(g^{-1}(y))} \mathrm{d}y$$

Now consider the change of variables $x = g^{-1}(y)$ (so that $\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}y} g^{-1}(y) \mathrm{d}y$, i.e. $g'(x)\mathrm{d}x = \mathrm{d}y$). We get

$$\int_{-\infty}^{\infty} y \cdot \frac{f(g^{-1}(y))}{g'(g^{-1}(y))} \mathrm{d}y = \int_{-\infty}^{\infty} g(x) \cdot \frac{f(x)}{g'(x)} g'(x) \mathrm{d}x = \int_{-\infty}^{\infty} g(x) \cdot f(x) \mathrm{d}x.$$