Statistics 394, Problem Set 1 Solutions

Wellner; 1/12/2000

1. K, 1.1, # 5, but assume that the dice are not fair; instead, if Y denotes the outcome of tossing a single die, assume that $P(Y=i)$ is proportional to i for $i=1,\ldots,6$. First, find these six probabilities. Then solve questions a(ii), a(iii), a(v), and b.

   Solution: Since $P(Y=i)=Ci$ for $i=1,\ldots,6$, we must have

$$1=\sum_{i=1}^{6}P(Y=i)=C(1+2+3+4+5+6)=C6\cdot 7/2=21\cdot C\,,$$

   and hence $C=1/21$.

   a(ii):

$$\begin{aligned} P(T=6) &= P(Y_1=1)P(Y_2=5)+P(Y_1=2)P(Y_2=4)+P(Y_1=3)P(Y_2=3)\\ &\qquad +\ P(Y_1=4)P(Y_2=2)+P(Y_1=5)P(Y_2=1)\\ &= \{1\cdot 5+2\cdot 4+3\cdot 3+4\cdot 2+5\cdot 1\}/(21)^2\\ &= 35/(21)^2\approx .0794\,. \end{aligned}$$

   a(iii):

$$\begin{aligned} P(\text{double}) &= P(Y_1=1,Y_2=1)+P(Y_1=2,Y_2=2)+P(Y_1=3,Y_2=3)+P(Y_1=4,Y_2=4)\\ &= P(Y_1=1)P(Y_2=1)+P(Y_1=2)P(Y_2=2)+P(Y_1=3)P(Y_2=3)\\ &\qquad +\ P(Y_1=4)P(Y_2=4)+P(Y_1=5)P(Y_2=5)+P(Y_1=6)P(Y_2=6)\\ &= \{1^2+2^2+3^2+4^2+5^2+6^2\}/(21)^2\\ &= 91/(21)^2\approx .2063\,. \end{aligned}$$

   a(v):

$$\begin{aligned} P(\text{at least one of } Y_1,Y_2>4) &= 1-P(Y_1\le 4,Y_2\le 4)\\ &= 1-P(Y_1\le 4)P(Y_2\le 4)=1-(10/21)^2\approx .7732\,. \end{aligned}$$

   b: With $X=Y_1+Y_2$, the events $[X=x]$ decompose as follows:

$$\begin{aligned} [X=2] &= [Y_1=1,Y_2=1]\\ [X=3] &= [Y_1=1,Y_2=2]\cup[Y_1=2,Y_2=1]\\ [X=4] &= [Y_1=1,Y_2=3]\cup[Y_1=2,Y_2=2]\cup[Y_1=3,Y_2=1]\\ [X=5] &= [Y_1=1,Y_2=4]\cup[Y_1=2,Y_2=3]\cup[Y_1=3,Y_2=2]\cup[Y_1=4,Y_2=1]\\ [X=6] &= [Y_1=1,Y_2=5]\cup[Y_1=2,Y_2=4]\cup[Y_1=3,Y_2=3]\\ &\qquad \cup[Y_1=4,Y_2=2]\cup[Y_1=5,Y_2=1]\\ [X=7] &= [Y_1=1,Y_2=6]\cup[Y_1=2,Y_2=5]\cup[Y_1=3,Y_2=4]\\ &\qquad \cup[Y_1=4,Y_2=3]\cup[Y_1=5,Y_2=2]\cup[Y_1=6,Y_2=1]\\ [X=8] &= [Y_1=2,Y_2=6]\cup[Y_1=3,Y_2=5]\cup[Y_1=4,Y_2=4]\\ &\qquad \cup[Y_1=5,Y_2=3]\cup[Y_1=6,Y_2=2]\\ [X=9] &= [Y_1=3,Y_2=6]\cup[Y_1=4,Y_2=5]\cup[Y_1=5,Y_2=4]\cup[Y_1=6,Y_2=3]\\ [X=10] &= [Y_1=4,Y_2=6]\cup[Y_1=5,Y_2=5]\cup[Y_1=6,Y_2=4]\\ [X=11] &= [Y_1=5,Y_2=6]\cup[Y_1=6,Y_2=5]\\ [X=12] &= [Y_1=6,Y_2=6]\ . \end{aligned}$$

Straightforward computation as in a(ii) above yields the following table:

k	2	3	4	5	6	7	8	9	10	11	12
$(21)^2 \cdot p_X(x)$	1	4	10	20	35	56	70	76	73	60	36
$p_X(x) \approx$	.002	.009	.023	.045	.079	.127	.159	.172	.166	.136	.082

2. K, 1.1, # 8, but do with *five* slips of paper numbered $1, 2, 3, 4, 5$.

Solution: Now there are two draws, without replacement, from a hat containing slips numbered $1, 2, 3, 4, 5$. Let $X_1 \equiv$the number drawn on the first draw, and $X_2 \equiv$ the number drawn on the second draw. Then we have:
a(i): $P(\text{slip 2 drawn}) = P([Y_1 = 2] \cup [Y_2 = 2]) = P(Y_1 = 2) + P(Y_2 = 2) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$.
a(ii):

$$\begin{aligned} P(\text{slip 3 drawn, slip 4 not drawn}) &= P(([Y_1 = 3] \cap [Y_2 \neq 4]) \cup ([Y_1 \neq 4] \cap [Y_2 = 3])) \\ &= P([Y_1 = 3] \cap [Y_2 \neq 4]) + P([Y_1 \neq 3 \text{ or } 4] \cap [Y_2 = 3]) \\ &= \frac{1}{5} \cdot \frac{3}{4} + \frac{3}{5} \cdot \frac{1}{4} \\ &= \frac{6}{20} . \end{aligned}$$

a(iii):

$$\begin{aligned} &P(\text{max of two numbers is not more than 3}) \\ &\quad = P(\max\{Y_1, Y_2\} \leq 3) = P(Y_1 \leq 3, Y_2 \leq 3) \\ &\quad = \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10} = \frac{\binom{3}{2}\binom{2}{0}}{\binom{5}{2}} . \end{aligned}$$

b: If $X = Y_1 + Y_2$, then X has possible values $3, \ldots, 9$, and we calculate, for example,

$$\begin{aligned} P(X = 6) &= P(Y_1 = 1, Y_2 = 5) + P(Y_1 = 2, Y_2 = 4) + P(Y_1 = 4, Y_2 = 2) + P(Y_1 = 5, Y_2 = 1) \\ &= \frac{1}{5}\frac{1}{4} + \frac{1}{5}\frac{1}{4} + \frac{1}{5}\frac{1}{4} + \frac{1}{5}\frac{1}{4} = 4/20 = 1/5 . \end{aligned}$$

Thus we find:

k	3	4	5	6	7	8	9
$10 \cdot p_X(x)$	1	1	2	2	2	1	1

c(i): $P(Y_1 = 3) = 1/5$, $P(Y_2 = 3) = 1/5$. To see the second probability, note that we can write it as

$$P(Y_2 = 3) = \sum_{j=1, j \neq 3}^{5} P(Y_2 = 3 | Y_1 = j) P(Y_1 = j) = \frac{1}{4} \cdot \frac{1}{5} + \frac{1}{4} \cdot \frac{1}{5} + \frac{1}{4} \cdot \frac{1}{5} + \frac{1}{4} \cdot \frac{1}{5} = \frac{1}{5} .$$

3. K, 1.2, # 5, but assume that $P(HEAD) = 2/3$, not $1/2$.

Solution: (a). The possible outcomes are $H, TH, TTH, TTTH, \ldots$; thus a general outcome for the experiment consists of $i-1$ tails (T), followed by a head (H). Assuming that the probability of a head on each toss of the coin is $2/3$, these various outcomes have corresponding probabilities $2/3$, $(1/3)(2/3) = 2/9$, $(1/3)^2(2/3)$, $(1/3)^3(2/3), \ldots$. Thus the general outcome of $i-1$ tails followed by a head has probability $(1/3)^{i-1}(2/3) = 2/3^i$, $i = 1, 2, 3, \ldots$. Note that these do sum to 1:

$$\sum_{i=1}^{\infty}(2/3)(1/3)^{i-1} = (2/3)\sum_{j=0}^{\infty}\frac{1}{3^j} = \frac{2/3}{1-1/3} = 1\,.$$

(b). Let $Y \equiv$ the number of tosses needed to get a head (H). Then by (a), $P(Y = i) = (1/3)^{i-1}(2/3)$ for $i = 1, 2, 3, \ldots$. Thus

$$\begin{aligned}
&P(\text{more than k tosses needed to get a head})\\
&= P(Y > k) = \sum_{i=k+1}^{\infty}(1/3)^{i-1}(2/3)\\
&= (1/3)^k(2/3) + (1/3)^{k+1}(2/3) + \cdots\\
&= (1/3)^k(2/3)\left\{1 + (1/3) + (1/3)^2 + \cdots\right\}\\
&= (1/3)^k(2/3)\frac{1}{1-1/3} = (1/3)^k\,.
\end{aligned}$$

Thus $P(Y > 2) = 1/9$, $P(Y > 3) = 1/27$, $P(Y > 4) = 1/81, \ldots$.
(c).

$$\begin{aligned}
P(Y \text{ is odd}) &= P([Y=1]\text{or}[Y=3]\text{or}[Y=5]\text{or}\ldots)\\
&= \sum_{j=1}^{\infty}(1/3)^{2j-2}(2/3)\\
&= (2/3)\sum_{m=0}^{\infty}(1/9)^m\\
&= (2/3)\frac{1}{1-1/9} = \frac{2}{3}\cdot\frac{9}{8} = \frac{3}{4}\,.
\end{aligned}$$

4. K, 1.2, # 19, but assume that only 10 people favor A while the other 1235 favor B.

Solution: (a) The probability of getting at most one supporter of A in a sample of size 7 without replacement, if there are only 10 people favoring A in the population, is:

$$\begin{aligned}
&\frac{\binom{10}{0}\binom{1235}{7}}{\binom{1245}{7}} + \frac{\binom{10}{1}\binom{1235}{6}}{\binom{1245}{7}}\\
&= \frac{1235}{1245}\cdot\frac{1234}{1244}\cdot\frac{1233}{1243}\cdot\frac{1232}{1242}\cdot\frac{1231}{1241}\cdot\frac{1230}{1240}\cdot\frac{1229}{1239}\\
&\qquad + \frac{70}{1239}\cdot\frac{1235}{1245}\cdot\frac{1234}{1244}\cdot\frac{1233}{1243}\cdot\frac{1232}{1242}\cdot\frac{1231}{1241}\cdot\frac{1230}{1240}\\
&\approx .94498 + .05382 = .9988\,.
\end{aligned}$$

(b) If the sampling is carried out with replacement, then the probability of at most one supporter of A in the sample is

$$\begin{aligned} &\binom{7}{0}\left(\frac{10}{1245}\right)^0\left(\frac{1235}{1245}\right)^7+\binom{7}{1}\left(\frac{10}{1245}\right)^1\left(\frac{1235}{1245}\right)^6 \\ = & \left(\frac{1235}{1245}\right)^7+7\left(\frac{10}{1245}\right)\left(\frac{1235}{1245}\right)^6 \\ \approx & .94511+.05357=.9987\,. \end{aligned}$$

Thus the probabilities under sampling with and without replacement differ only by about .0001.

5. K, 1.3, # 7: justify your answers.

Solution: Yes: c and g. Not possible: e and f. Possible to find a constant to make them densities: a,b,d. Here's why:
(a) $x^4 \geq 0$ for $0 \leq x \leq 1$ and $\int_0^1 x^4 dx = (1/5)x^5|_0^1 = (1/5)$, so $f(x) = 5x^4 1_{[0,1]}(x)$ is a density function.
(b) $x^4 \geq 0$ for $0 \leq x \leq 2$, and $\int_0^2 x^4 dx = (1/5)x^5|_0^2 = 32/5$, so $f(x) = (5/32)x^4 1_{[0,2]}(x)$ is a density.
(c) $x^{-2} \geq 0$ for $x \geq 1$, and $\int_1^\infty x^{-2} dx = (-1)^{-1}x^{-1}|_1^\infty = 1$. Thus $f(x) = x^{-2}1_{[1,\infty)}(x)$ is a density function.
(d) $\sin(x) \geq 0$ for $0 \leq x \leq \pi$ and

$$\int_0^\pi \sin(x)dx = -\cos(x)|_0^\pi = -(-1-1) = 2\,.$$

Thus $f(x) = (1/2)\sin(x)1_{[0,\pi]}(x)$ is a density function.
(e) $\cos(x) < 0$ for $\pi/2 < x \leq \pi$, so it is not possible to normalize this function to be a density function.
(f) $1/x \geq 0$ for $x > 0$, but

$$\int_0^\infty \frac{1}{x} dx = \log(x)|_1^\infty = \infty - (-\infty) = \infty\,.$$

Thus this f cannot be renormalized to be a density function.
(g) $x\exp(-x^2/2) \geq 0$ for $x \geq 0$, and, by changing to the new variable $y = x^2/2$,

$$\int_0^\infty x\exp(-x^2/2)dx = \int_0^\infty e^{-y}dy = -e^{-y}|_0^\infty = -(0-1) = 1\,.$$

Thus this function f is a density function on $[0,\infty)$.

6. Let $\Omega = [0,1)$ with the uniform distribution as in example 1.3.1. Define the random variable Z by $Z(\omega) = -\log(\omega)$. In part e, name the distribution.

Solution: (a) $P(Z \leq 2) = P(-\log(\omega) \leq 2) = P(\log(\omega) \geq -2) = P(\omega \geq e^{-2}) = 1 - e^{-2}$.

(b) $P(Z > 3) = P(-\log(\omega) > 3) = P(\log(\omega) < -3) = P(\omega < e^{-3}) = e^{-3}$.
(c)

$$\begin{aligned} P(3 \leq Z \leq 4) = &= P(3 \leq -\log(\omega) \leq 4) == P(-3 \geq \log(\omega) \geq -4) \\ &= P(e^{-3} \geq \omega \geq e^{-4}) = e^{-3} - e^{-4} . \end{aligned}$$

(d) $P(a \leq Z \leq b) = e^{-a} - e^{-b} = 1 - e^{-b} - (1 - e^{-a})$.
(e) Thus Z has an exponential distribution with parameter 1.

7. K, 1.3, # 14: omit parts a,b,c,d. Instead find the probability of the event that the random point lies within 1/2 unit of the center *and* lies in the first quadrant of the disk.

   Solution: Since the circle of radius 1 has area $\pi(1)^2 = \pi$, and the set of points in the first quadrant which are within 1/2 of the center of the disk has area $(1/4)\pi(1/2)^2 = \pi/16$, the probability of getting a point in the first quadrant within 1/2 of the center of the disk is $(\pi/16)/\pi = 1/16$.