Name solutions

Section

Instructions: read each problem carefully and provide solutions in the space below the prompt or on the reverse side of the page. You should provide as much justification and detail in calculation as needed to clearly display your thought process, but need not show or justify every step. If you use results from class in your solution, you can simply write, "by a theorem/problem/example from class...". If you are unable to obtain a complete solution, a solution sketch may receive partial credit if it reflects a clear understanding of the problem and a well-reasoned approach. Please feel free to ask any clarifying questions about the problems as they arise. Good luck!

1. Let (X, Y) be distributed according to the PDF $f(x, y) = e^{-y}$ for $0 < x < y < \infty$. Find the covariance and correlation of X and Y.

$f_X(x) = \int_x^\infty e^{-y}\,dy = e^{-x}, \quad 0 < x < \infty \implies X \sim \Gamma(1,1)$

$f_Y(y) = \int_0^y e^{-y}\,dx = y e^{-y}, \quad 0 < y < \infty \implies Y \sim \Gamma(2,1)$

so:

$\mathbb{E}X = 1 \qquad \text{var}\,X = 1$

$\mathbb{E}Y = 2 \qquad \text{var}\,Y = 2$

last:

$\mathbb{E}XY = \int_0^\infty \int_0^y xy e^{-y}\,dx\,dy = \frac{1}{2}\int_0^\infty y^3 e^{-y}\,dy = \frac{\Gamma(4)}{2}\int_0^\infty \underbrace{\frac{1}{\Gamma(4)1^4} y^{4-1} e^{-y}}_{\Gamma(4,1)\text{ PDF}}\,dy = \frac{\Gamma(4)}{2} = \frac{3!}{2} = 3$

so:

$\text{cov}(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y = 3 - 2 \cdot 1 = 1$

$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}\,X\,\text{var}\,Y}} = \frac{1}{\sqrt{2}}$

2. About 5% of adults in the continental U.S. have coronary artery disease (CAD). From demographic data, the probabilities that a randomly selected individual lives in the northeast, southeast, south, midwest, or west, are approximately:

NE	S	MW	W
0.2	0.2	0.2	0.4

a) Fill in the table below with a joint probability distribution in which region and CAD rate are independent. Is this distribution unique? Explain.

	NE	S	MW	W
CAD	0.01	0.01	0.01	0.02
healthy	0.19	0.19	0.19	0.38

for independence, must have P(status, region) = P(status) P(region) so with the marginals fixed this is the only possible solution

a) Is it possible that the conditional probability $P(CAD|W)$ is 0.2? Why or why not?

no: $P(CAD|W)P(W) = P(CAD \cap W) \leq P(CAD)$

so $P(CAD|W) \leq \frac{P(CAD)}{P(W)} = \frac{0.05}{0.4} = 0.125$

$P(CAD|W) = 0.2$ would imply $P(CAD) \geq 0.08$

3. Zero-inflated probability models modify existing distributions by adding excess probability mass at zero. These are useful for modeling data with many zeroes. Let $Y \sim \text{Poisson}(\lambda)$ and $Z \sim \text{Bernoulli}(p)$. let X be a random variable such that:

$$P(X = x|Z = 1) = \begin{cases} 1 & , x = 0 \\ 0 & , x \neq 0 \end{cases}$$
$$P(X = x|Z = 0) = P(Y = x)$$

The distribution of X is called a "zero-inflated Poisson". Determine the PMF of X and find its mean.

$$P(X=x) = P(X=x|Z=0)P(Z=0) + P(X=x|Z=1)P(Z=1)$$
$$= (1-p)P(Y=x) + p \cdot P(X=x|Z=1)$$
$$= \begin{cases} (1-p)e^{-\lambda} + p \cdot 1 & , x=0 \\ (1-p)\dfrac{\lambda^x e^{-\lambda}}{x!} + p \cdot 0 & , x = 1, 2, \ldots \end{cases}$$
$$= \begin{cases} p + (1-p)e^{-\lambda} & , x=0 \\ (1-p)\dfrac{\lambda^x e^{-\lambda}}{x!} & , x = 1, 2, 3, \ldots \end{cases}$$

$$\mathbb{E}X = \sum_{x=0}^{\infty} x P(X=x)$$
$$= \sum_{x=1}^{\infty} x P(X=x)$$
$$= (1-p)\sum_{x=1}^{\infty} x P(Y=x)$$
$$= (1-p)\sum_{x=0}^{\infty} x P(Y=x)$$
$$= (1-p)\mathbb{E}Y$$
$$= (1-p)\lambda$$

4. Let X_1, X_2 be independent Gaussian random variables with means and variances μ_1, μ_2 and σ_1^2, σ_2^2, respectively. Find the MGF of $X_1 + X_2$ and use this to determine its distribution.

let $Y = X_1 + X_2$

$$m_{X_1}(t) = \exp\left\{\mu_1 t + \tfrac{1}{2} t^2 \sigma_1^2\right\}$$

$$m_{X_2}(t) = \exp\left\{\mu_1 t + \tfrac{1}{2} t^2 \sigma_2^2\right\}$$

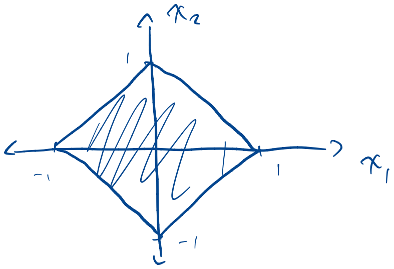
by independence

$$m_Y(t) = \mathbb{E} e^{tY} = \mathbb{E} e^{t(X_1 + X_2)} = \mathbb{E} e^{tX_1} e^{tX_2} = \mathbb{E} e^{tX_1} \, \mathbb{E} e^{tX_2} = m_{X_1}(t)\, m_{X_2}(t)$$

$$= \exp\left\{ \underbrace{(\mu_1 + \mu_2)}_{\text{new } \mu} t + \tfrac{1}{2} t^2 \underbrace{(\sigma_1^2 + \sigma_2^2)}_{\text{new } \sigma^2} \right\}$$

so $Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

5. Let X_1, X_2 be jointly uniform on the set $|x_1|+|x_2| \le 1$. Sketch the support set, determine the joint PDF, and then find the conditional distribution of X_1 given $X_2 = x_2$. Based on your answer, are X_1, X_2 dependent?



support set

$S = \{(x_1, x_2) : |x_1| + |x_2| \le 1\}$

uniform means $f(x_1, x_2) = c$

must have $\iint_S f(x_1, x_2)\, dx_1\, dx_2 = 1$

but $\iint_S c\, dx\, dy = c \cdot \text{area}(S) = c \cdot 2$

so $f(x_1, x_2) = \frac{1}{2}, \quad |x_1| + |x_2| \le 1$

marginal of X_2:

$$f_2(x_2) = \int_{|x_2|-1}^{1-|x_2|} \tfrac{1}{2}\, dx_1 = \tfrac{1}{2} \cdot (2 - 2|x_2|) = 1 - |x_2|, \quad -1 \le x_2 \le 1$$

$|x_1| + |x_2| \le 1 \iff |x_1| \le 1 - |x_2| \iff -(1-|x_2|) \le x_1 \le (1 - |x_2|)$

conditional of $X_1 | X_2$:

$$f_{1|2}(x_1) = \frac{f(x_1, x_2)}{f(x_2)} = \frac{1}{2(1-|x_2|)}, \quad |x_2| - 1 \le x_1 \le 1 - |x_2| \implies X_1 | X_2 \sim \text{uniform}(|x_2|-1, 1-|x_2|)$$

$X_1 \not\perp X_2$ because (any of the following are valid answers):

- $f_{1|2}(x_1)$ depends on x_2
- support of $f_{1|2}$ depends on x_2
- $f_{1|2}(x_1) \neq f_1(x_1)$
  $(2(1-|x_2|))^{-1} \neq 1 - |x_1|$
- $f(x_1, x_2) \neq f_1(x_1) f_2(x_2)$
  $\frac{1}{2} \neq (1 - |x_1|)(1 - |x_2|)$

Name	PDF/PMF	Support	Parameters	Mean	Variance	MGF
Discrete uniform	$\frac{1}{n}$	$\{a_1, \ldots, a_n\}$	none	$\bar{a}$	$\frac{1}{n}\sum_{i=1}^{n}(a_i - \bar{a})^2$	$\frac{1}{n}\sum_{i=1}^{n} e^{ta_i}$
Bernoulli	$p^x(1-p)^{1-x}$	$\{0, 1\}$	$p \in (0, 1)$	p	$p(1-p)$	$1 - p + pe^t$
Geometric	$(1-p)^x p$	$\mathbb{N}$	$p \in (0, 1)$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$	$\frac{p}{1-(1-p)e^t}$, $t < -\log(1-p)$
Binomial	$\binom{n}{p} p^x (1-p)^{n-x}$	$\{0, 1, \ldots, n\}$	$n \in \mathbb{Z}^+$ $p \in (0, 1)$	np	$np(1-p)$	$(1 - p + pe^t)^n$
Negative binomial	$\binom{r+x-1}{x} p^r (1-p)^x$	$\mathbb{N}$	$r \in \mathbb{Z}^+$ $p \in (0, 1)$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{p}{1-(1-p)e^t}\right)^r$, $t < -\log(1-p)$
Poisson	$\frac{\lambda^x e^{-\lambda}}{x!}$	$\mathbb{N}$	$\lambda > 0$	λ	λ	$\exp\{-\lambda(1-e^t)\}$
Continuous uniform	$\frac{1}{b-a}$	(a, b)	$-\infty < a < b < \infty$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	$\begin{cases} \frac{e^{ta}-e^{tb}}{t(b-a)} & , t \neq 0 \\ 1 & , t = 0 \end{cases}$
Gaussian	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$	$\mathbb{R}$	$\mu \in \mathbb{R}, \sigma > 0$	μ	σ^2	$\exp\left\{\mu t + \frac{1}{2}t^2\sigma^2\right\}$
Gamma	$\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}$	$x > 0$	$\alpha > 0, \beta > 0$	$\alpha\beta$	$\alpha\beta^2$	$(1-\beta t)^{-\alpha}$, $t < \frac{1}{\beta}$