Question 5.1 If $\mathbb{E}X^r = \infty$ show that

$$\mathbb{E}X^r = r\int_0^\infty x^{r-1}P(x \geq x)dx$$

Solution: We will integrate by parts to show the result. Let $u = rP(X \geq x) = r - rF(X)$. Then $du = -rf(x)dx$. Moreover, let $dv = x^{r-1}dx$. Then $v = \frac{1}{r}x^r$. Then we have

$$r\int_0^\infty x^{r-1}P(x \geq x)dx = (1 - F(x))x^r\Big|_0^\infty + \frac{1}{r}\int_0^\infty x^r f(x)dx$$

$$= P(X \geq x)x^r\Big|_{x\to\infty} + \frac{1}{r}\mathbb{E}(X^r) \geq \frac{1}{r}\mathbb{E}(X^r) = \infty$$

Thus, equality holds if $\mathbb{E}(X^r) = \infty$

Question 5.2 Let A_1, A_2 be nontrivial events (i.e. $P(A_1), P(A_2) > 0$). Show that A_1,A_2 disjoint does not imply A_1,A_2 are independent. Show that A_1, A_2 independent does not imply A_1, A_2 disjoint.

Solution:

Assume that $A_1 \cap A_2 = \emptyset$. Then $P(A_1 \cap A_2) = P(\emptyset) < P(A_1)P(A_2)$. Therefore $P(A_1 \cap A_2) \neq P(A_1)P(A_2)$ so A_1 and A_2 are not independent.

Assume that A_2 and A_2 are independent. Then $P(A_1 \cap A_2) = P(A_1)P(A_2) > 0$. Thus $A_1 \cap A_2 \neq \emptyset$.

Question 5.3 Give an example that shows that pairwise independence does not imply mutual independence.

Solution: Consider two die rolls. Let A be the event that the first roll is even, B is the event that the second roll is even, and C be the event that the sum of the two rolls is even. Then $P(A) = P(B) = P(C) = \frac{1}{2}$. Moreover we have $P(A \cap B) = \frac{1}{4} = P(A)P(B)$. Now notice for the sum of the two rolls to be even, either both rolls must be even or both rolls must be odd. So the events $A \cap C$ and $B \cap C$ require both rolls be even. Hence $P(A \cap C) = P(B \cap C) = \frac{1}{4} = P(A)P(C) = P(B)P(C)$. Thus we see these events are pairwise independent. Now consider $P(A \cap B \cap C)$. Notice the event $A \cap B \subset C$ so $A \cap B \cap C = A \cap B$. Thus $P(A \cap B \cap C) = P(A \cap B) = 1/4$. But notice that $P(A)P(B)P(C) = \frac{1}{8}$. Thus

$$P(A \cap B \cap C) \neq P(A)P(B)P(C)$$

and we see that pairwise independence does not imply mutual independence.

Question 5.4 Let $\{X_k\}_k^n$ be a sequence of mean zero random variables that are weakly stationary. Define $R_{ij} = r(i-j) = Cov(X_i, X_j)$. Using this notation, show

$$Var\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i=1}^{n}\sum_{j=1}^{n} r(i-j) = nr(0) + 2A_n$$

where $A_n = \sum_{k=1}^{n-1}(n-k)r(k)$

Solution: First let $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$. Then the $Var(\mathbf{X}) = \mathbf{\Sigma}$ can be written as $\mathbf{\Sigma}_{ij} = r(i-j)$. Let $\mathbf{1}_n = (1, 1, \dots, 1)^T$ be a vector of n ones. Then our problem reduces to

$$Var(\mathbf{1}_n^T\mathbf{X}) = \mathbf{1}_n^T Var(\mathbf{X})\mathbf{1}_n = \mathbf{1}_n^T\mathbf{\Sigma}\mathbf{1_n}$$

Now notice that this is just the sum over all elements in $\mathbf{\Sigma}$. Now, all n elements on the main diagonal are $r(0)$. The diagonal above (and below) the main diagonal have $n-1$ elements and are all $r(1)$. In general, for the $k = 1, 2, \dots, n-1$ diagonal above/below the main diagonal there are $n-k$ elements all of the form $r(k)$. This yields

$$\mathbf{1_n^T}\mathbf{\Sigma}\mathbf{1_n} = nr(0) + 2\sum_{k=1}^{n-1}(n-k)r(k)$$

Question 5.5 Suppose $X_1 \sim$Gamma(p_1, λ) and $X_2 \sim$Gamma(p_2, λ). Find the distribution of $X_1 + X_2$.

Solution: We use the convolution formula to find the distribution of $X_1 + X_2$.

$$\begin{aligned}
f_{X_1+X_2}(x) &= \int_{-\infty}^{\infty} f_1(x-y)f_2(y)dy \\
&= \int_{-\infty}^{\infty} \Big(\frac{1}{\Gamma(p_1)}(x-y)^{p_1-1}\lambda^{p_1}e^{-\lambda(x-y)}\Big)\Big(\frac{1}{\Gamma(p_2)}y^{p_2-1}\lambda^{p_2}e^{-\lambda y}\Big)dy \\
&= \int_{-\infty}^{\infty} \frac{1}{\Gamma(p_1)\Gamma(p_2)}(x-y)^{p_1-1}y^{p_2-1}\lambda^{p_1+p_2}e^{-\lambda x}dy \\
&= \frac{x^{p_1+p_2-2}\lambda^{p_1+p_2}e^{-\lambda x}}{\Gamma(p_1+p_2)}\int_{-\infty}^{\infty}\frac{\Gamma(p_1+p_2)}{\Gamma(p_1)\Gamma(p_2)}\frac{(x-y)^{p_1-1}y^{p_2-1}}{x^{p_1-1}x^{p_2-1}}dy \\
&= \frac{x^{p_1+p_2-2}\lambda^{p_1+p_2}e^{-\lambda x}}{\Gamma(p_1+p_2)}\int_{-\infty}^{\infty}\frac{\Gamma(p_1+p_2)}{\Gamma(p_1)\Gamma(p_2)}(1-y/x)^{p_1-1}(y/x)^{p_2-1}dy
\end{aligned}$$

After completing a change of variables $u = y/x$ we see that the integrand is the density of a Beta distribution. Therefore, it integrates to one and can write

$$= \frac{x^{p_1+p_2-2}\lambda^{p_1+p_2}e^{-\lambda x}}{\Gamma(p_1+p_2)}\int_{-\infty}^{\infty}\frac{\Gamma(p_1+p_2)}{\Gamma(p_1)\Gamma(p_2)}(1-u)^{p_1-1}(u)^{p_2-1}xdu = \frac{x^{p_1+p_2-1}\lambda^{p_1+p_2}e^{-\lambda x}}{\Gamma(p_1+p_2)}$$

which we recognize as a density of a Gamma with parameters (p_1+p_2, λ). Thus $X_1 + X_2 \sim$Gamma(p_1+p_2, λ).

Question 5.6 Show that if X_1 and X_2 are independent then

$$P(X_1 + X_2 = n) = \sum_{k=1}^{n} P(X_1 = n-k)P(X_2 = k)$$

Using this, show that if $Y_1 \sim$Pois(λ_1) and $Y_2 \sim$Pois(λ_2) are independent, then $Y_1 + Y_2 \sim$Pois($\lambda_1 + \lambda_2$)

Solution: Let $\mathbf{X}$ be the support of $X_1 + X_2$. Using the fact that X_1 and X_2 are independent and the law of total probability, we have

$$\begin{aligned}
P(X_1 + X_2 = n) &\overset{LTP}{=} \sum_{k\in\mathbf{X}} P(X_1 + X_2 = n | X_2 = k)P(X_2 = k)\\
&= \sum_{k\in\mathbf{X}} P(X_1 = n-k | X_2 = k)P(X_2 = k)\\
&\overset{ind.}{=} \sum_{k\in\mathbf{X}} P(X_1 = n-k)P(X_2 = k)
\end{aligned}$$

Now consider the probability mass function of $Y_1 + Y_2$.

$$\begin{aligned}
P(Y_1 + Y_2 = n) &= \sum_{k=0}^{\infty} P(Y_1 = n-k)P(Y_2 = k)\\
&= \sum_{k=0}^{\infty} \frac{e^{-\lambda_1}\lambda_1^{n-k}}{(n-k)!}\frac{e^{-\lambda_2}\lambda_2^{k}}{(k)!}\\
&= e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{\infty} \frac{\lambda_1^{n-k}\lambda_2^{k}}{(n-k)!k!}\\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{i=0}^{\infty} \frac{n!}{(n-k)!n!}\lambda_1^{n-k}\lambda_2^{k}\\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!} \sum_{k=0}^{\infty} \binom{n}{k} (\frac{\lambda_1}{\lambda_1+\lambda_2})^{n-k}(\frac{\lambda_2}{\lambda_1+\lambda_2})^{k}\\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!} \sum_{k=0}^{\infty} \binom{n}{k} (1-\frac{\lambda_2}{\lambda_1+\lambda_2})^{n-k}(\frac{\lambda_2}{\lambda_1+\lambda_2})^{k}
\end{aligned}$$

We recognize the sum as the probability mass function of a binomial random variable with probability $\frac{\lambda_2}{\lambda_1+\lambda_2}$. Thus we see

$$P(Y_1 + Y_2) = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!}$$

which is the density of a Poisson random variable with parameter $\lambda_1 + \lambda_2$. Hence $Y_1 + Y_2 \sim$Pois($\lambda_1 + \lambda_2$)

Question 5.7 If $(X, Y) \sim BVN(\mu, \mathbf{\Sigma})$ where $\mu = (\mu_X, \mu_Y)^T$ and

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

Find the conditional distribution $f_{X|Y}(x|y)$.

Solution: First recall that the marginal distribution of a bivariate random variable is normal. We will use the definition of marginal distribution $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(u)}$.

$$\begin{aligned}
f_{X|Y}(x|y) &= \frac{\frac{1}{2\pi\sigma_Y\sigma_X\sqrt{1-\rho^2}}\exp\Big\{-\frac{1}{2(1-\rho^2)}\Big((\frac{y-\mu_Y}{\sigma_Y})^2 - 2\rho(\frac{y-\mu_Y}{\sigma_Y})(\frac{x-\mu_X}{\sigma_X}) + (\frac{x-\mu_X}{\sigma_X})^2\Big)\Big\}}{\frac{1}{\sqrt{2\pi}\sigma_Y}\exp\{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\}} \\
&= \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}}\exp\Big\{-\frac{1}{2(1-\rho^2)}\Big(\rho^2(\frac{y-\mu_Y}{\sigma_Y})^2 - 2\rho(\frac{y-\mu_Y}{\sigma_Y})(\frac{x-\mu_X}{\sigma_X}) + (\frac{x-\mu_X}{\sigma_X})^2\Big)\Big\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}}\exp\Big\{-\frac{1}{2(1-\rho^2)}\Big(\rho(\frac{y-\mu_Y}{\sigma_Y}) - (\frac{x-\mu_X}{\sigma_X})\Big)^2\Big\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}}\exp\Big\{-\frac{1}{2\sigma_X^2(1-\rho^2)}\Big(\rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y) - (x-\mu_X)\Big)^2\Big\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}}\exp\Big\{-\frac{1}{2\sigma_X^2(1-\rho^2)}\Big(x-\mu_X-\rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y)\Big)^2\Big\}
\end{aligned}$$

We recognize this as density of a normal distribution. Specifically

$$X|Y \sim N\left(\mu_x - \rho(\sigma_X/\sigma_Y)(y-\mu_Y), \sigma_X^2(1-\rho^2)\right)$$

Question 5.8 Suppose there are $n+m$ Bernoulli trials with probability p. After we observe n successes, find the posterior distribution of p if the prior distribution of p is uniform.

Solution: By definition we have

$$f_{p|n}(p|n) = \frac{P(N=n)f_P(p)}{P(N=n)} = \frac{\binom{n+m}{n}p^n(1-p)^m}{\int_0^1 \binom{n+m}{n}p^n(1-p)^m dp}$$

Focusing on the denominator, we have

$$\begin{aligned}
\int_0^1 \binom{n+m}{n}p^n(1-p)^m dp &= \frac{(n+m)!}{n!m!}\int_0^1 p^n(1-p)^m dp \\
&= \frac{\Gamma(n+m+1)}{\Gamma(n+1)\Gamma(m+1)}\int_0^1 p^{(n+1)-1}(1-p)^{(m+1)-1}dp \\
&= \frac{1}{n+m+1}\left[\frac{\Gamma(n+m+2)}{\Gamma(n+1)\Gamma(m+1)}\int_0^1 p^{(n+1)-1}(1-p)^{(m+1)-1}dp\right]
\end{aligned}$$

We recognize the bracketed term as the density of a Beta with parameters $(n+1, m+1)$. Therefore, it integrates to 1 we are left with $\frac{1}{n+m+1}$. Plugging this into our original form we have

$$f_{P|N}(p|n) = (n+m+1)\binom{n+m}{n}p^n(1-p)^m = \frac{\Gamma(n+m_2)}{\Gamma(n+1)\Gamma(m+1)}p^{(n+1)-1}(1-p)^{(m+1)-1}$$

Thus the posterior is a Beta random variable with parameters $(n+1, m+1)$.