Lsningar till tentamen i Matematisk statistik och diskret matematik D (MVE055/MSG810)
Den 25 oktober 2016.

1. Since $P(A) = P(A|B) = 0.9$, therefore A and B are independent and $P(A \cap B) = P(A)P(B)$. Then, we have:

$$\begin{aligned} P(A|B^c) &= \frac{P(B^c|A)P(A)}{P(B^c)} \\ &= \frac{(1 - P(B|A))P(A)}{1 - P(B)} \\ &= \frac{(1 - \frac{P(A\cap B)}{P(A)})P(A)}{1 - P(B)} \\ &= \frac{(1 - P(B))P(A)}{1 - P(B)} \\ &= P(A) = 0.9 \end{aligned}$$

One can conclude that if A and B are independent, then A and B^c are independent too.

2. (a) $m_X(t) = E(e^{tX})$, which is the generating function of the sequence of the moments of the random variable X.

Let $X \sim N(\mu, \sigma^2)$, then

$$\begin{aligned} m_X(t) = E(e^{tX}) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} e^{tx} dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{-\mu^2}{2\sigma^2}} e^{-\frac{1}{2\sigma^2}(x^2-2\mu x-2\sigma^2 tx)} dx \\ &= e^{-\frac{-\mu^2}{2\sigma^2}+\frac{1}{2\sigma^2}(\mu+\sigma^2 t)^2} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-(\mu+\sigma^2 t))^2} dx \\ &= e^{\mu t + \frac{t^2}{2}\sigma^2} \end{aligned}$$

where the final equality follows from the fact that the expression under the integral is the $N(\mu + \sigma^2 t, \sigma^2)$ probability density function which integrates to unity.

(b)Let $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, 2, ..., n$ and X_i's are independent random variables. Then,

$$\begin{aligned}
m_{\bar{X}}(t) &= E(e^{t\bar{X}}) \\
&= E(e^{\frac{t}{n}\sum_{i=1}^{n} X_i}) \\
&= E(e^{\frac{t}{n}X_1}.e^{\frac{t}{n}X_2}....e^{\frac{t}{n}X_n}) \\
&= E(e^{\frac{t}{n}X_1}).E(e^{\frac{t}{n}X_2})....E(e^{\frac{t}{n}X_n}) \\
&= m_{X_1}(t/n).m_{X_2}(t/n)....m_{X_n}(t/n) \\
&= e^{\mu_1 t/n+\frac{t^2}{2n^2}\sigma_1^2}.e^{\mu_2 t/n+\frac{t^2}{2n^2}\sigma_2^2}....e^{\mu_n t/n+\frac{t^2}{2n^2}\sigma_{n1}^2} \\
&= \prod_{i=1}^{n} e^{\mu_i t/n+\frac{t^2}{2n^2}\sigma_i^2} \\
&= e^{\sum_{i=1}^{n}(\mu_i t/n+\frac{t^2}{2n^2}\sigma_i^2)}.
\end{aligned}$$

3. $X, Y \sim U[0,1]$ and they are independent. Define $U = min(X,Y)$ and $V = max(X,Y)$. Then, $cov(U,V) = E(UV) - E(U)E(V)$.
$F_X(x) = x, 0 \leq x \leq 1$ and $F_Y(y) = y, 0 \leq y \leq 1$.

The cdf of U is:

$$\begin{aligned}
F_U(u) = P(U \leq u) &= P(min(X,Y) \leq u) \\
&= 1 - P(min(X,Y) \geq u) \\
&= 1 - P(X \geq u)P(Y \geq u) \\
&= 1 - (1 - F_X(u))^2 \\
&= 1 - (1-u)^2.
\end{aligned}$$

and the pdf of U is:

$$f_U(u) = F_U'(u) = 2(1-u), 0 \leq u \leq 1$$

Then, the expectation of U is:

$$E(U) = \int_0^1 u f_U(u) du = \int_0^1 2u(1-u) du = 1/3.$$

The cdf of V is:

$$\begin{aligned}
F_V(v) = P(V \leq v) &= P(max(X,Y) \leq v) \\
&= P(X \leq v)P(Y \leq v) \\
&= F_X(v)^2 \\
&= v^2.
\end{aligned}$$

and the pdf of V is:

$$f_V(v) = F'_V(v) = 2v, 0 \leq v \leq 1$$

Then, the expectation of V is:

$$E(V) = \int_0^1 v f_V(v) dv = \int_0^1 2v^2 dv = 2/3.$$

$$\begin{aligned}
cov(U,V) &= E(UV) - E(U)E(V) \\
&= E(XY) - E(U)E(V) \\
&= E(X)E(Y) - E(U)E(V) \\
&= \frac{1}{2} \times \frac{1}{2} - \frac{1}{3} \times \frac{2}{3} = \frac{1}{36}.
\end{aligned}$$

4. (a) $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$ and they are independent. To find the distribution of $X+Y$, we use moment generating function:

$$\begin{aligned}
m_{X+Y}(t) &= E(e^{t(X+Y)}) \\
&= E(e^{tX}).E(e^{tY}) \\
&= m_X(t).m_Y(t) \\
&= e^{\lambda_1(e^t-1)}.e^{\lambda_2(e^t-1)} \\
&= e^{(\lambda_1+\lambda_2)(e^t-1)}
\end{aligned}$$

Therefore, $X+Y \sim Poi(\lambda_1+\lambda_2)$.

(b)

$$\begin{aligned}
P(X=x|X+Y=n) &= \frac{P(X+Y=n|X=x)P(X=x)}{P(X+Y=n)} \\
&= \frac{P(Y=n-x|X=x)P(X=x)}{P(X+Y=n)} \\
&= \frac{P(Y=n-x)P(X=x)}{P(X+Y=n)} \\
&= \frac{\frac{e^{-\lambda_2}\lambda_2^{n-x}}{(n-x)!} \cdot \frac{e^{-\lambda_1}\lambda_1^x}{x!}}{\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!}} \\
&= \binom{n}{x}(\frac{\lambda_1}{\lambda_1+\lambda_2})^x(\frac{\lambda_2}{\lambda_1+\lambda_2})^{n-x}
\end{aligned}$$

Therefore, $X|X+Y \sim Bin(n, \frac{\lambda_1}{\lambda_1+\lambda_2})$.

5. If X and Y be independent, then $E\left[g(X)h(Y)\right] = E\left[g(X)\right]E\left[h(Y)\right]$. If $f(x)$ be the pdf of (X), then $E\left[g(X)\right] = \int_x f_X(x)g(x)dx$.

In this question, $X, Y \sim U[0,1]$, and $f_X(x) = f_Y(y) = 1$ for $0 \le x \le 1, 0 \le y \le 1$. Therefore:

(a) $E[XY] = E[X]E[Y]$,

(b) $E[X/Y] = E[X]E[1/Y]$ and $E[1/Y] = \int_0^1 \frac{1}{y}dy$,

(c) $E[\log(XY)] = E[log(X)] + E[log(Y)] = \int_0^1 \log x dx. \int_0^1 \log y dy$.

6. There are two samples $X_1, ..., X_{n_A} \sim N(\mu_A, \sigma_A^2)$ and $Y_1, ..., Y_{n_B} \sim N(\mu_B, \sigma_B^2)$, where $n_A = 7$ and $n_B = 5$, and $\sigma_A^2 = \sigma_B^2 = \sigma^2$(unknown). The 90% confidence intervals for the difference in mean vitamin E content between the two types is:

$$\mu_A - \mu_B \in (\bar{X}_A - \bar{X}_B \pm t_{1-\alpha/2}(n_A + n_B - 2) S_P \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}),$$

where $\bar{X}_A$ and $\bar{X}_B$ are the sample means, $S_P^2 = \frac{(n_A-1)S_A^2 + (n_B-1)S_B^2}{n_A+n_B-2}$ is the pooled variance and $t_{0.05}(10) = 1.812$.

7. $\bar{x} = 4.3$, $\bar{y} = 12.9$, $\sum x_i^2 = 123.25$, $\sum y_i^2 = 839.09$ and $\sum x_i y_i = 262.75$.

$$\begin{aligned} \hat{\beta}_1 &= \frac{\sum x_i y_i - 5\bar{x}\bar{y}}{\sum x_i^2 - 5\bar{x}^2} = -0.47 \\ \hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x} = 14.94 \end{aligned}$$

The fitted regression line is:

$$\hat{y} = 14.94 - 0.47x.$$

To do the test $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$, we use the following test statistic:

$$T = \frac{\hat{\beta}_1}{S/\sqrt{S_{xx}}}$$

where $S_{xx} = \sum x_i^2 - 5\bar{x}^2$ and $S = \frac{SSE}{n-2}$. Therefore, we have:

$$T = \frac{-0.47}{\sqrt{0.0397/30.8}} = -13.2$$

and,

$$P - value = 2.min[P(T > -13.2), P(T < -13.2)] \approx 0.0001$$

which is much less than $\alpha = 0.05$ and we reject H_0.