Solution to Bivariate Normal Problem

---

a) The marginal distribution of Y is not affected by the correlation between X and Y; in particular, $Y \sim N(7,\ 9)$. So, no matter what the correlation between X and Y, we have

$$P(Y > 10) = P\left(\frac{Y-7}{3} > \frac{10-7}{3}\right) = P\left(Z > 1\right) = 1 - \Phi(1) = 0.1587.$$

b) The information given yields $Y|X = x \sim N(7 + 3\rho(x-2),\ 9(1-\rho^2))$ for any value of x.

So, $Y|X = 2 \sim N(7,\ 9(1-\rho^2))$. In particular, $Y|X = 2 \sim N(7,\ 9)$, $Y|X = 2 \sim N(7,\ 6.75)$ and $Y|X = 2 \sim N(7,\ 1.71)$, when $\rho = 0,\ 0.5$ and 0.9, respectively. Calculations as in a) then yield $P(Y > 10|X = 2) = 0.1587,\ 0.1241$ and 0.0109 when $\rho = 0,\ 0.5$ and 0.9, respectively.

Note how the distribution of $Y|X = 2$ changes as ρ changes: the mean is not affected (because $\mu_X = 2$) but the stronger the correlation, the smaller the variance of this conditional distribution.

c) Similarly as in b), $Y|X = 4 \sim N(7 + 6\rho,\ 9(1-\rho^2))$. So, we obtain $Y|X = 4 \sim N(7,\ 9)$, $Y|X = 4 \sim N(10,\ 6.75)$ and $Y|X = 4 \sim N(12.4,\ 1.71)$, when $\rho = 0,\ 0.5$ and 0.9, respectively. Calculations as in a) then yield $P(Y > 10|X = 4) = 0.1587,\ 0.5000$ and 0.9668 when $\rho = 0,\ 0.5$ and 0.9, respectively.

Note that the impact of the correlation on the variance of the distribution of $Y|X = x$ is the same at every value of x but, in this case, the mean of this conditional distribution also changes as ρ changes (because the value of x under consideration differs from $\mu_X = 2$ and the slope of the regression line increases as the correlation increases).

d) Here we need to use the basic result that if X and Y are bivariately normally distributed then any linear combination of X and Y is normally distributed. (As indicated in class, this is easy to show using moment generating functions.) So, we simply need to evaluate the mean and the variance of the linear combination of interest. But $E(4X - Y) = 4E(X) - E(Y) = 4(2) - 7 = 1$ and $\text{Var}(4X - Y) = 16\text{Var}(X) + \text{Var}(Y) + 2(4)(-1)\text{Cov}(X, Y) = 16(1) + 9 - 8\rho(1)(3) = 25 - 24\rho$.

So, $4X - Y \sim N(1,\ 25 - 24\rho)$. In particular, $4X - Y \sim N(1,\ 25)$, $N(1,\ 13)$ and $N(1,\ 3.4)$, when $\rho = 0,\ 0.5$ and 0.9, respectively. Calculations as in a) then yield $P(4X - Y > 11) = 0.0228,\ 0.0028$ (roughly) and 0 (essentially), when $\rho = 0,\ 0.5$ and 0.9, respectively.

e) Exactly as in d). $E(3X + 2Y) = 3E(X) + 2E(Y) = 3(2) + 2(7) = 20$ and $\text{Var}(3X + 2Y) = 9\text{Var}(X) + 4\text{Var}(Y) + 2(3)(2)\text{Cov}(X, Y) = 9(1) + 4(9) + 12\rho(1)(3) = 45 + 36\rho$.

So, $3X + 2Y \sim N(20,\ 45 + 36\rho)$. In particular, $3X + 2Y \sim N(20,\ 45)$, $N(20,\ 63)$ and $N(20,\ 77.4)$, when $\rho = 0,\ 0.5$ and 0.9, respectively. Calculations as in a) then yield $P(3X + 2Y > 30) = 0.0680,\ 0.1038$ and 0.1278, when $\rho = 0,\ 0.5$ and 0.9, respectively.