- 1. (i)  $X + 2Y \sim N(\mu_X + 2\mu_Y, \sigma_X^2 + 4\sigma_Y^2)$ , where  $\mu_X = -1$ ,  $\sigma_X^2 = 2$ ,  $\mu_Y = 1$ , and  $\sigma_Y^2 = 3$ . Therefore,  $X + 2Y \sim N(-1 + 2 * 1, 2 + 4 * 3) = N(1, 14)$ . So, the correct answer is [B.] N(1, 14).
  - (ii) As previously discussed,  $\frac{(X+Y)^2}{(X-Y)^2}$  follows an F-distribution with 1 and 1 degrees of freedom. So, the correct answer is [A.]  $F_{1,1}$ .
  - (iii) The following are random variables:
    - The sample mean: This changes with each sample.
    - The smallest value in the sample: This changes with each sample.
    - The estimated variance of the sample mean: This depends on the sample. So, the correct choices are [E.], [F.], and [G.].

2. (a) 
$$f_X(x) = \int_0^1 \frac{2}{5} (2x+3y) dy = \frac{2}{5} \left( 2xy + \frac{3}{2}y^2 \right) \Big|_0^1 = \frac{4}{5}x + \frac{3}{5}, \quad 0 \le x \le 1$$

(b) 
$$f_{Y|x}(y) = \frac{\frac{2}{5}(2x+3y)}{\frac{4}{5}x+\frac{3}{5}} = \frac{4x+6y}{4x+3}, \quad 0 \le y \le 1$$

(c) 
$$f_{Y|\frac{1}{2}}(y) = \frac{1}{5}(2+6y)$$
 
$$P(1/4 \le Y \le 3/4) = \int_{1/4}^{3/4} \frac{1}{5}(2+6y)dy = 10/20 = 1/2$$

(d) 
$$\mathbb{E}(Y \mid X = x) = \int_0^1 y \cdot f_{Y|x}(y) dy$$
$$= \int_0^1 y \cdot \frac{4x + 6y}{4x + 3} dy = \frac{4x/2 + 2}{4x + 3} = \frac{2x + 2}{4x + 3}.$$

3. This can be shown using the definition of the F-distribution with m and n degrees of freedom and the t-distribution with n degrees of freedom: if X is a standard normal random variable, Y a  $\chi^2$  random variable with n degrees of freedom, and X and Y are independent, then

$$T = \frac{X}{\sqrt{\frac{Y}{n}}}$$

follows a t-distribution with n degrees of freedom. Therefore, T squared is

$$T^2 = \frac{\frac{X^2}{1}}{\frac{Y}{n}},$$

and we know that  $X^2$  follows a  $\chi^2$ -distribution with 1 degree of freedom. Therefore, from the definition of an F distribution, it's clear that  $T^2$  follows an F-distribution with 1 and n degrees of freedom, as was to be shown.

**4.** (25 points)

(a) The sample mean  $\overline{X}$  is calculated as follows:

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i = \frac{1}{9} (105 + 107 + 110 + 104 + 92 + 90 + 87 + 116 + 89) = 100$$

(b) To calculate the unbiased estimate of the population variance, we first calculate

$$\hat{\sigma_n}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X})^2 = \frac{900}{9} = 100,$$

then adjust it by a factor to get the unbiased estimate. The unbiased estimate of the population variance  $\sigma^2$  is then given by:

$$\hat{\sigma}_{n, \text{ unbiased}}^2 = \frac{(N-1)n}{N(n-1)} \cdot \hat{\sigma}_n^2 = \frac{1360 \cdot 9 - 9}{1360 \cdot 9 - 1360} \cdot 100 \approx 112.42$$

The unbiased estimate of the variance of the sample mean  $Var(\overline{X})$  is given by:

$$S_{\overline{X}}^2 = \frac{\hat{\sigma}_{n, \text{ unbiased}}^2}{n} \left( 1 - \frac{n-1}{N-1} \right) = 12.49 \cdot \frac{1351}{1359} \approx 12.42$$

- (c) To calculate a 95% confidence interval for the population mean, we use the formula  $\overline{X} \pm z_{\alpha/2} \cdot S_{\overline{X}}$ , where  $S_{\overline{X}}$  is the estimated standard error of the sample mean, given by  $\sqrt{S_{\overline{X}}^2}$ . For a 95% confidence interval,  $z_{2.5\%} \approx 1.96$ . Therefore, the 95% confidence interval for the population mean is  $100 \pm 1.96 \cdot 3.52 = (93.09, 106.91)$ .
- 5. For a Poisson distribution, the mean  $\mu$  is equal to the variance. Given that the sample mean  $\overline{X} = 3.4$ , it is a point estimate of  $\mu$ .

When the sample size is large (n > 30) is a common rule of thumb), the Central Limit Theorem can be used to approximate the sampling distribution of the sample mean as a normal distribution, even if the underlying distribution (in this case, Poisson) is not normal.

The standard error (SE) is given by the standard deviation of the distribution divided by the square root of the sample size. For a Poisson distribution, the standard deviation is  $\sqrt{\mu}$ , so the standard error can be estimated as

$$S_{\overline{X}} = \sqrt{\overline{X}/n} = \sqrt{\frac{3.4}{200}} = 0.1304$$

With  $Z_{5\%} = 1.645$ , the 90% confidence interval for  $\mu$  is approximately  $3.4 \pm Z_{5\%} \cdot \sqrt{\frac{\overline{X}}{n}}$ , which is [3.186, 3.614].