Câu 1. (L.O.1, L.O.2.1, L.O.2.3)

(a) (1.5đ = 0.5 +1)

- Gọi X là số hành khách đến trạm này trong 10 phút, thì $X \sim Poison(4)$.
- Xác suất để có nhiều nhất 3 hành khách đến trạm này trong 10 phút:

$$P(X \leq 3) = \sum_{k=0}^{k=3} P(X = k) = 0.4335$$

(b) (1.5đ = 0.5+1)

- Gọi Y(phút) là thời gian mà xe buýt phải chờ để đón thêm ít nhất một hành khách nữa, thì $Y \sim Exp(0.4)$.
- $\mathbb{E}(Y) = \dfrac{1}{0.4} = 2.5$ (phút), $S(Y) = \sqrt{\mathbb{V}ar(Y)} = \sqrt{\dfrac{1}{0.4^2}} = 2.5$ (phút), $\mathbb{E}(Y) + 2S(Y) = 7.5$ (phút).
- Giả sử xe đã chờ 5 phút mà không có hành khách nào, xác suất để xe phải chờ thêm ít nhất 5 phút nữa là:
  Cách 1:

$$P(Y \geq 10|Y > 5) = P(Y \geq 5 + 5|Y > 5) = P(Y \geq 5) \text{ (Tính chất không nhớ)}$$
$$= 1 - P(Y < 5) = 1 - \int_0^5 0.4e^{-0.4t}dt = e^{-2}$$

Cách 2:

$$P(Y \geq 10|Y > 5) = \frac{P(Y > 10)}{P(Y > 5)} = \frac{1 - \int_0^{10} 0.4e^{-0.4t}dt}{1 - \int_0^5 0.4e^{-0.4t}dt} = e^{-2}$$

Câu 2. (L.O.1, L.O.2.1, L.O.2.3)

(a) (1đ)
Theo công thức bảng B:

- Tính $S_{xx} = 1485.71406$, $S_{xy} = 9678.571$, $\overline{y} = 397.1429$, $\overline{x} = 31.4286$.
- Tính $\hat{\beta}_1 = \dfrac{S_{xy}}{S_{xx}} \approx 6.5144$, $\hat{\beta}_0 = \overline{y} - \hat{\beta}_1\overline{x} \approx 192.4038$.
- Tại $X_0 = 35$, tính $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 X_0 \approx 420.4087$ (kg).

Theo công thức bảng A:

- Tính $\overline{xy} = 13864.29$, $\overline{y} = 397.1429$, $\overline{x} = 31.4286$, $\hat{S}_x^2 = 212.2449$.
- Tính Tính $B \approx 6.5144$, $A \approx 192.4038$.
- Tại $X_0 = 35$, tính $\hat{Y}_0 = A + BX_0 \approx 420.4087$ (kg).

(b) (0.5đ)
Theo công thức bảng B:

- Tính $STT = \sum_i (y_i - \overline{y})^2 = 63792.86$, $R^2_{XY} = \beta_1^2 \dfrac{S_{xx}}{STT} \approx 0.9884$, $\Rightarrow$ hệ số tương quan mẫu: $R_{XY} = 0.9942$.

- X và Y có mối quan hệ tuyến tính mạnh.

Theo công thức bảng A:

- Tính $\hat{S}_y \approx 95.4635$, $r_{xy} = \dfrac{\overline{xy} - \overline{x}\overline{y}}{\hat{S}_x \hat{S}_y} \approx 0.9942$
- X và Y có mối quan hệ tuyến tính mạnh.

Câu 3. (L.O.1, L.O.2.1, L.O.2.3) (2đ)

- Giả thuyết H0: Màu tóc **không** có sự ảnh hưởng đến sức chịu đau của phụ nữ:
  $\tau_1 = \tau_2 = \tau_3 = 0$
  vs H1: Màu tóc **có** sự ảnh hưởng đến sức chịu đau của phụ nữ: $\exists i = 1, 2, 3,\ \tau_i \neq 0$.
- Tính $SSB = 654$, $SSW = 654.25$, $SST = 1308.25$, $MSB = 327$, $MSW = 72.7$.
- Tính $df(SSB) = 2$, $df(SSW) = 9$, $df(SST) = 11$, $F = \dfrac{MSB}{MSW} \approx 4.49828$.
- Tính $f_{0.05,2,9} = 4.256$.
- $F > f_{0.05,2,9}$ nên ta bác bỏ giả thuyết H0. Dữ liệu khảo sát cho thấy có sự ảnh hưởng của màu tóc đến sức chịu đau của phụ nữ.

Câu 4. (L.O.1, L.O.2.1, L.O.2.3)

(a) (1đ)

- Tính $\hat{P} = \dfrac{10}{200} = 0.05$, $z_{0.05} = 1.96$.
- Khoảng tin cậy cho tỷ lệ các chi tiết do máy A sản xuất là:

$$\hat{P} - z_{0.05}\sqrt{\frac{\hat{P}(1-\hat{P})}{200}} \leq P \leq \hat{P} + z_{0.05}\sqrt{\frac{\hat{P}(1-\hat{P})}{200}} \Leftrightarrow 0.0198 \leq P \leq 0.0802$$

(b) (1đ)

- Giả thuyết H0: $\mu_A = 25$ vs H1: $\mu_A < 25$.
- Tính $\overline{X} = 24.59$, $s_A = 0.9146$, $t_0 = \dfrac{\overline{X} - \mu_0}{s_A/\sqrt{n_A}} = -1.4175$.
- Tính $t^9_{0.05} = 1.8331$.
- $t_0 > -t^9_{0.05}$ nên chưa đủ cơ sở để bác bỏ giả thuyết $H0$. Với 95% độ tin cậy ta chưa đủ cơ sở để nói rằng chiều dài trung bình của các chi tiết sản xuất bởi máy A là ngắn hơn 25 cm.

(c) (1.5đ)

- Giả thuyết H0: $\mu_A = \mu_B$ vs H1: $\mu_A \neq \mu_B$.
- Tính $z_0 = \dfrac{\overline{X} - \overline{Y}}{\sqrt{\dfrac{\sigma_A^2}{n_A} + \dfrac{\sigma_B^2}{n_B}}} = \dfrac{24.59 - 25}{\sqrt{\dfrac{0.9^2}{10} + \dfrac{0.9^2}{190}}} = -1.4041$.
- $z_{0.005} = 2.576$.
- $|z_0| < z_{0.005}$ nên chưa đủ cơ sở để bác bỏ giả thuyết $H0$. Với 99% độ tin cậy ta chưa đủ cơ sở để nói rằng các chi tiết do hai máy sản xuất có chiều dài trung bình khác nhau.