Question 1.

(a) With $\lambda = 0.25$, we have

$$P(Y \leq 1) = F_Y(1) = 1 - e^{-0.25(1)} \approx 0.221.$$

(b) We set $F_Y(\phi_{0.5}) = 0.5$ and solve for $\phi_{0.5}$. We have

$$\begin{aligned}
1 - e^{-0.25\phi_{0.5}} &\overset{\text{set}}{=} 0.5 \\
\Longrightarrow e^{-0.25\phi_{0.5}} &= 0.5 \\
\Longrightarrow -0.25\phi_{0.5} &= \ln 0.5 \quad \Longrightarrow \quad \phi_{0.5} = -\frac{1}{0.25}\ln 0.5 \approx 2.77.
\end{aligned}$$

(c) Define

$$\begin{aligned}
A_1 &= \{\text{component 1 survives at least 10 years}\} \\
A_2 &= \{\text{component 2 survives at least 10 years}\} \\
A_3 &= \{\text{component 3 survives at least 10 years}\} \\
A_4 &= \{\text{component 4 survives at least 10 years}\}.
\end{aligned}$$

The reliability of the parallel system is $P(A_1 \cup A_2 \cup A_3 \cup A_4)$. Note that

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = 1 - P(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3 \cap \overline{A}_4) = 1 - P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3)P(\overline{A}_4).$$

The probability component 1 fails before 10 years is $P(\overline{A}_1) = F_Y(10)$, with $\lambda = 0.25$. This is $1 - e^{-0.25(10)} \approx 0.918$. This the same failure probability for component 2. The probability component 3 fails before 10 years is $P(\overline{A}_3) = F_Y(10)$, with $\lambda = 0.5$. This is $1 - e^{-0.5(10)} \approx 0.993$. This the same failure probability for component 4. Therefore, the reliability is

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) \approx 1 - (0.918)(0.918)(0.993)(0.993) \approx 0.169.$$

Question 2.

(a) In this situation, the three Bernoulli trial assumptions are:

1. Each block is either defective or it is not.
2. The status of each block is independent of other blocks' statuses.
3. The probability of being defective, $p = 0.07$, is the same for all blocks.

(b) We calculate

$$P(Y \leq 2) = P(Y = 0) + P(Y = 1) + P(Y = 2).$$

Use the binomial pmf to calculate each probability:

$$P(Y = 0) = \binom{20}{0}(0.07)^0(0.93)^{20} \approx 0.234$$

$$P(Y = 1) = \binom{20}{1}(0.07)^1(0.93)^{19} \approx 0.353$$

$$P(Y = 2) = \binom{20}{2}(0.07)^2(0.93)^{18} \approx 0.252.$$

Therefore,

$$P(Y \leq 2) \approx 0.234 + 0.353 + 0.252 = 0.839.$$

(c) The expected cost is

$$E(C) = E(50 + 4Y + 2Y^2) = 50 + 4E(Y) + 2E(Y^2).$$

We have $E(Y) = np = 20(0.07) = 1.4$. Also,

$$E(Y^2) = \text{var}(Y) + [E(Y)]^2 = 20(0.7)(0.93) + (1.4)^2 \approx 3.262.$$

Therefore,

$$E(C) = 50 + 4(1.4) + 2(3.262) \approx 62.12 \text{ dollars}.$$

Question 3.
(a) From the Central Limit Theorem, the sampling distribution of the sample mean $\overline{Y}$ is approximately normal; specifically,

$$\overline{Y} \sim \mathcal{AN}\left(\mu, \frac{\sigma^2}{n}\right),$$

where μ denotes the population mean solid count (per liter of water) for the river and σ^2 denotes the population variance solid count (per liter of water) for the river. The sample size $n = 40$.
(b) The sample mean $\overline{y} = 31.2$ is an estimate of μ, defined above. The sample variance $s^2 = 24.8$ is an estimate of σ^2, defined above.
(c) The standard error of the sample mean $\overline{Y}$ is

$$\text{se}(\overline{Y}) = \sqrt{\text{var}(\overline{Y})} = \sqrt{\frac{\sigma^2}{40}} = \frac{\sigma}{\sqrt{40}}.$$

An estimate of this quantity is

$$\frac{s}{\sqrt{40}} = \frac{\sqrt{24.8}}{\sqrt{40}} \approx 0.787.$$

Question 4.
(a) The qq-plot calculated under a Weibull assumption looks very linear. Therefore, the Weibull distribution looks to be a good population model for the ignition times.
(b) We are 95 percent confident that the population mean ignition time μ is between 4.61 and 5.82 seconds.
(c) The confidence interval in part (b) is a statement about the population mean ignition time. It does not tell us anything about individual materials' ignition times.
(d) This confidence interval is only useful if the underlying population distribution is normal, and this interval is not robust to departures from this assumption. However, we concluded in part (a) that a Weibull distribution was a good model for the population. Therefore, one might exercise caution in interpreting this interval.

Question 5.
(a) If we perform an analysis for this part's reliability under an incorrect population model assumption, then all of our resulting calculations could be misleading. For example, our estimate of the population hazard function could be drastically wrong. Probabilities of part failure and

quantile estimates could be incorrect. The end result is that our analysis and understanding of the part's reliability could be compromised.
(b) A 95 percent confidence interval for the Weibull shape parameter β is

$$1.188 \pm 1.96(0.350) \implies (0.502, 1.874).$$

We are 95 percent confident that the (population level) Weibull shape parameter β is between 0.502 and 1.874.
(c) Your co-worker is making a statement that is not supported by the analysis. The estimate of β is larger than one, but this is only the estimate. The confidence interval in part (b) includes "1" and values that are less than 1. Recall that if $\beta = 1$, then the population distribution is exponential (which corresponds to random failures). If $\beta < 1$, then the population of parts is not getting weaker over time; it is actually getting stronger over time. Your co-worker doesn't understand the difference between an estimate calculated from sample and the population level model.

Question 6.
(a) We are 95 percent confident that the population mean difference $\mu_1 - \mu_2$ is between -399.6 and -37.8 grams. Because this interval includes only negative values, this suggests that the population mean birth weight of babies born to mothers who use drugs is lower than the population mean birth weight of babies born to mothers who do not use drugs.
(b) There is no way this could be done ethically. To use a matched pairs design, you would have to observe a baby's birth weight under the "no drugs" condition and one under the "drugs" condition. We could not require a mother to use drugs.
(c) The sample proportions are

$$\begin{aligned}\widehat{p}_1 &= \frac{8}{65} \approx 0.123 \\ \widehat{p}_2 &= \frac{37}{550} \approx 0.067.\end{aligned}$$

A 95 percent confidence interval for $p_1 - p_2$, the difference of the population proportions, is

$$(0.123 - 0.067) \pm 1.96\sqrt{\frac{0.123(0.877)}{65} + \frac{0.067(0.933)}{550}} \implies (-0.027, 0.138).$$

Because this interval includes "0," we cannot conclude (at the 95 percent confidence level) that the population proportion of NEC-positive babies is different between the two groups of mothers (i.e., those who use drugs and those who do not).

Question 7.
(a) There is strong evidence that energy use and temperature are linearly related in the population. Note that the p-value for the test of

$$H_0 : \beta_1 = 0$$
versus
$$H_1 : \beta_1 \neq 0$$

is 0.0000177, which is incredibly small. This would be a significant finding at any reasonable level of significance (e.g., 0.05).
(b) In the simple linear regression model, 48.7 percent of the variation in the energy use data is explained by temperature. This leaves 51.3 percent of the variation in the energy use data not explained by the simple linear regression model.
(c) The residual plot looks very random in its appearance. This plot does not reveal any model inadequacies (i.e., it does not suggest that the simple linear regression model is inadequate).
(d) Because he wants to estimate a population mean, he wants to use a confidence interval. A prediction interval would only be used if he wants to make an inferential statement about an individual bird's energy usage.

Question 8.
(a) There is insufficient evidence to assert that rayon whiteness and acid temperature are linearly related in the population (after adjusting for the effects of the other variables). The large p-value (0.442) would not lead us to conclude that $\beta_1 \neq 0$ at any reasonable level of significance.
(b) The hypotheses are

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$$
and
$$H_1\text{: at least one of the } \beta_j\text{'s is nonzero.}$$

In words, H_0 means that none of the independent variables are important in describing rayon whiteness in the population. The alternative H_1 means that at least one of the independent variables is important in describing rayon whiteness in the population.
(c) The assumptions are

- $E(\epsilon_i) = 0$, for $i = 1, 2, ..., n$
- $\text{var}(\epsilon_i) = \sigma^2$, for $i = 1, 2, ..., n$, that is, the variance is constant
- the random variables ϵ_i are independent
- the random variables ϵ_i are normally distributed.

Question 9.
(a) The researcher's belief is not supported by the analysis. First of all, the interaction plot is nearly parallel. Secondly, the p-value for the test of interaction in the population is 0.173. This would not be declared as a significant result at any reasonable level of significance (e.g., $\alpha = 0.05$).
(b) Both the main effects of moisture and temperature are significant in the population. The corresponding p-values for the main effects are very small.
(c) The treatment sum of squares would be

$$\text{SS}_{trt} = 60.500 + 43.245 + 0.845 = 104.59.$$

(d) The treatment mean squares would be $\text{SS}_{trt}/3 \approx 34.86$. The residual mean squares is 0.307 (i.e., it does not change). Therefore,

$$F = \frac{\text{MS}_{trt}}{\text{MS}_{res}} \approx \frac{34.86}{0.307} \approx 113.56.$$