ST1100 SOLUTIONS TO TEST 2

[WITH THE QUESTIONS]

1. If f is the density function of a continuous random variable X, what must $\int_{-\infty}^{\infty} f(x)dx$ equal?

(A) 0 **(B)** 0.2 **(C)** 0.5 **(D)** 1 **(E)** π **(F)** ∞

SOLUTION: The answer is **(D)**.

Details: See the boxed information on *page 13* of *Some Probability Lecture Notes.* The total area under the density function of a continuous random variable is the probability that X lies between $-\infty$ and ∞ and this must be 1 as it is certain that X will lie between $-\infty$ and ∞.

2. Suppose that a large conference room at a certain company can be reserved for no more than 4 hours. Both long and short conferences occur quite often. In fact, it can be assumed that the length X of a conference has a uniform distribution on the interval $(0, 4)$. Thus the density function of X is

$$f(x) = \begin{cases} \frac{1}{4}, & 0 < x < 4 \\ 0, & \text{elsewhere} \end{cases}$$

What is the probability that a random conference will last at least 3 hours?

(A) 0.1 **(B)** 0.2 **(C)** 0.25 **(D)** 0.4 **(E)** 0.5 **(F)** 0.55

SOLUTION: The answer is **(C)**. *Details:* This is part (b) of *Example 1* on pages *187-188* of our *online textbook* entitled *Probability & Statistics for Engineers & Scientists. 9th Ed.* As you know, one way of accessing this book by going to the library web site and typing into the search box *Probability & Statistics for Engineers & Scientists*. The second link you see will have an ‘Online resource’ link which gives you access to the entire book.(Note that there is a slider at the top of the pages to move forward!)
Note: If you take out on loan the hard-copy of this book from the library, the present problem is on pages 171-172).
The solution to the present problem is given in the text and note that of course you can avoid the integration used there by instead using geometry... just get the area of a rectangle that has base length 1 [corresponding to the interval (3,4)] and height 1/4.

3. Refer to Question 2 above. Given that a conference will last at most two hours, what is the probability that it will last at most one hour? That is, among all conferences in the room that last at most two hours, what proportion last at most one hour?

(A) 0.1 **(B)** 0.2 **(C)** 0.25 **(D)** 0.4 **(E)** 0.5 **(F)** 0.55

SOLUTION: The answer is **(E)**.
Details: Following a conditional probability argument repeatedly performed in the course resources, here we have

$$P(X < 1 \mid X < 2) = \frac{P(X < 1 \cap X < 2)}{P(X < 2)} \underset{\text{why?}}{=} \frac{P(X < 1)}{P(X < 2)} =$$

$$\frac{\text{area of rectangle that has base (0,1) and height } \frac{1}{4}}{\text{area of rectangle that has base (0,2) and height } \frac{1}{4}} = \frac{1/4}{1/2} = \frac{1}{2}.$$

Of course since a uniform distribution has constant density, we were able to avoid using calculus.

If you insist on using calculus, you'd calculate $\frac{P(X<1)}{P(X<2)} = \frac{\int_{-\infty}^{1} f(x)dx}{\int_{-\infty}^{2} f(x)dx}$ where

$f(x) = \begin{cases} 1/4, 0 < x < 4 \\ 0,\ \text{elsewhere.} \end{cases}$.

This is $\frac{P(X<1)}{P(X<2)} = \frac{\int_0^1 1/4\,dx}{\int_0^2 1/4\,dx} = \frac{\frac{1}{4}x\Big|_0^1}{\frac{1}{4}x\Big|_0^2} = \frac{1/4}{1/2} = \frac{1}{2}$, as above

4. Suppose the time X in years till a random washing machine requires servicing has the exponential distribution with mean 4 years. Thus the density function of X is

$$f(x) = \begin{cases} \frac{1}{4}e^{-\frac{x}{4}}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

If a launderette buys five such machines, what is a numerical expression (in terms of e) for the probability that

exactly four of them will not require servicing for at least 6 years (that is, that exactly four will last more than

6 years without needing servicing)?

(A) $5e^{-6}(1-e^{-\frac{3}{2}})$ **(B)** $(1-e^{-\frac{3}{2}})^5$ **(C)** $4(1-e^{-\frac{3}{2}})^4$
(D) $e^{-\frac{3}{2}}(1-e^{-\frac{3}{2}})$ **(E)** $0.5e^{-1}$ **(F)** $0.05e$

SOLUTION: The answer is **(A)**.

Details: We first calculate the probability that one random machine will not require servicing for at least 6 years. This is $P(X \geq 6)$. This is solved in *Example 21* on *pages 215-216* of the *online textbook* (*example 6.21* on *pages 199-200* if using the hard copy in the library), and is similar to the solution to *part (a)* of the *example* on *page 15* of *Some Probability Lecture Notes.* We have

$P(X \geq 6) = \int_6^\infty \frac{1}{4}e^{-x/4}dx = -e^{-x/4}\Big|_6^\infty = 0 - (e^{-6/4}) = e^{-3/2}$

Now among the five machines purchased by the launderette, let W = number of machines that will not need servicing till after 6 years. We require $P(W \geq 4)$. Notice that W has the binomial distribution with parameters $n = 5$ and $p = P$(a random machine lasts more than 6 years)= $e^{-3/2}$. Then following a similar example in the lectures, we obtain

$P(W=4) = \binom{5}{4}(e^{-3/2})^4(1-e^{-3/2})^1 = 5e^{-6}(1-e^{-3/2})$

5. Let X denote the decay time of some radioactive particle and assume that X has the exponential density

$$f(x) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

Suppose θ is such that $P(X \geq 0.01) = 1/2$. What is a numerical expression for the number t that satisfies

$P(X \geq t) = 0.9$?

(A) $e^{\frac{0.9}{2}}$ **(B)** $e^{\frac{2}{0.9}}$ **(C)** $\frac{-\ln(2)}{\ln(0.9)}$ **(D)** $\frac{-\ln(0.9)}{100\ln(2)}$ **(E)** $-\ln(0.9)$ **(F)** $100\ln(2)$

SOLUTION: The answer is **(D)**.

Details: Since θ satisfies $P(X \geq 0.01) = 1/2$, we have $\int_{0.01}^\infty \frac{1}{\theta}e^{-x/\theta}dx = 1/2$. Hence $-e^{-x/\theta}\Big|_{0.01}^\infty = 1/2$ and so $\exp(\frac{-0.01}{\theta}) = 1/2$ leading to $\theta = \frac{0.01}{\ln(2)} = \frac{1}{100\ln(2)}$ *Note:* $\exp(y)$ is of course another

way of writing e^y
But from $P(X \geq t) = 0.9$, we have $\int_t^\infty \frac{1}{\theta} e^{-x/\theta} dx = 0.9$, i.e. $-e^{-x/\theta}\Big|_t^\infty = 0.9$, i.e. $e^{-t/\theta} = 0.9$. Hence $t = -\theta \ln(0.9)$, and since we had found $\theta = \frac{1}{100\ln(2)}$, we thus has
$t = -\frac{1}{100\ln(2)} \ln(0.9) = \frac{-\ln(0.9)}{100\ln(2)}$

6. The population of weekly expenditures on alcohol by students has a normal distribution with mean $\mu = €25$ and standard deviation $\sigma = €2.50$. If 100 students will be selected at random, what approximately is the expected number who will have expenditures above €30?

(A) 50 **(B)** 15.87 **(C)** 2.28 **(D)** 0.13 **(E)** 5

SOLUTION: The answer is **(C)**.

*Details:*This is *Problem 2* on *Problem Set 2* (with a slight change in figures) and the details are rather similar to those in the solution to the *example [Heights]* on *page 23* of *Statistics Lecture Notes*)
We first get the proportion of all students whose expenditures exceed 30.
Let X = weekly expenditure of a random student. We are given that $X \backsim N(\mu, \sigma^2)$ with $\mu = 25$ and $\sigma^2 = (2.5)^2$.
The proportion of the population with weekly expenditures above 25 is then
$P(X > 25) \underset{\text{standardizing}}{=} P\left(\frac{X-\mu}{\sigma} > \frac{25-\mu}{\sigma}\right) \underset{\text{where } Z \backsim N(0,1)}{=} P(Z > \frac{30-25}{2.5}) = P(Z > 2) \underset{\text{from } N(0,1) \text{ tables}}{=} 0.0228$.
In a random sample of 100 people, we'd thus expect 100×0.0228 =$\underline{2.28}$ to have weekly expenditures exceeding 30.

Side Note: If you are unclear on why we multiplied 0.0228 by 100, recall the formula *np* for the mean of a
binomial random variable that has parameters *n* and *p*. See e.g. *Statistics Lecture Notes page 11*.

7. Suppose that a population of marks of students has a normal distribution with mean $\mu = 65$ and standard deviation $\sigma = 15$. Let $a = P$(mark of a random student falls between 35 and 80) and let $b = P$(the mean mark for a random sample of 225 students will be between 63 and 66). Then it is true that

(A) $a = 0.6826$ and $b = 0.9544$ **(B)** $a = 0.9544$ and $b = 0.9544$
(C) $a = 0.9544$ and $b = 0.6826$ **(D)** $a = 0.8185$ and $b = 0.9772$
(E) $a = 0.8185$ and $b = 0.8185$ **(F)** $a = 0.8185$ and $b = 0.0013$

SOLUTION: The answer is **(E)**.
Details: $a = P(35 \leq X \leq 80) \underset{\text{standardizing}}{=} P(\frac{35-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{80-\mu}{\sigma}) =$
$P(\frac{35-65}{15} \leq Z \leq \frac{80-65}{15}) = P(-2 < Z < 1) = 1 - \{P(Z < -2) + P(Z > 1)\} =$
$1 - \{P(Z > 2) + P(Z > 1)\} = 1 - \{0.0228 + 0.1587\} = 1 - 0.1815 = 0.8185$
Also
$b = P(63 \leq \overline{X} \leq 66) \underset{\text{standardizing}}{=} P(\frac{63-\mu}{\frac{\sigma}{\sqrt{n}}} \leq \frac{X-\mu}{\frac{\sigma}{\sqrt{n}}} \leq \frac{66-\mu}{\frac{\sigma}{\sqrt{n}}}) = P\left(\frac{63-65}{\frac{15}{\sqrt{225}}} < Z < \frac{66-65}{\frac{\sigma}{\sqrt{n}}}\right) =$
$P(-2 < Z < 1) = 0.8185$) as in our previous calculation.

8. Student *A* will take a random sample of size 100 from an infinite population. Student *B* will take a random sample of size *n* from the same population. How large should *n* be if it is desired that the standard deviation of the mean $\overline{X}_B$ of *B*'s sample be *three times smaller* than the standard deviation of the mean $\overline{X}_A$ of *A*'s sample?

A) 10 **B)** 30 **C)** 50 **D)** 150 **E)** 450 **F)** 900

SOLUTION: The answer is **(F)**.
Details: This is *Problem 10* on *Problem Set 2* (with a change in one figure).
We want $\frac{1}{3}\frac{\sigma}{\sqrt{100}} = \frac{\sigma}{\sqrt{n}}$. Hence $\frac{1}{\sqrt{900}} = \frac{1}{\sqrt{n}}$ so $n = 900$

9. The diameters of bolts from a certain machine are normally distributed with mean $\mu = 3$ and standard deviation $\sigma = 0.001$. Let X be the diameter of a randomly selected bolt and let $\overline{X}$ be the mean diameter of a random sample of 100 bolts. Consider the following five statements.

(1) $P(|X-3| < 0.1) = P(|\overline{X}-3| < 0.01)$ **(2)** $P(3-0.001 < \mu < 3+0.001) = 0.5$

(3) $\overline{X}$ is a parameter **(4)** μ and σ are statistics **(5)** $E(X) = E(\overline{X})$ and $Var(\overline{X}) = \dfrac{(0.001)^2}{100}$

Then it is true that
(A) None of statements (1)–(5) is true
(B) Exactly one of statements (1)–(5) are true
(C) Exactly two of statements (1)–(5) are true
(D) Exactly three of statements (1)–(5) are true
(E) Exactly four of statements (1)–(5) are true
(F) All five of statements (1)–(5) are true

SOLUTION: The answer is **(C)**.
Details: Note that (1), (2),(3) and (4) were discussed at length in the solution to *Question 9* on *Problem Sheet 2*. Regarding (5), see *Lecture Notes in Statistics*, from *bottom of page 22 to page 26*. The first main point in (5) is that both X and $\overline{X}$ are centered at μ. That is, $E(X) = \mu$ and $E(\overline{X}) = \mu$; that is the average of all the individuals in the population is μ and the average of all the sample means (from all the various samples of size n) is also μ. (We then say that each of X and $\overline{X}$ are *unbiased* estimators of μ).
The second point to note is that while the variance of the individuals in the population is σ^2, that is $Var(X) = \sigma^2$, the means $\overline{X}$ of all the various samples of size n are not as spread out (taking means reduces variability and this is VERY important!) In fact $Var(\overline{X}) = \frac{\sigma^2}{n}$.

Useful note: In light of the veracity of both statements in (5), ask yourself this very important question: If we did not know μ (the mean of the population) would you prefer to use X (one random measurement taken from the population) or the mean $\overline{X}$ of a large random sample of size n? Of course you'd prefer the latter, because while X and $\overline{X}$ both have expected value equal to μ, the variance of $\overline{X}$ is much smaller than the variance of X, so we'd be more likely to get a value of $\overline{X}$ closer to μ than of getting a value of X close to μ. See also the *example* on *pages 25-36* of *Statistics Lecture Notes* where this important point is emphasised. If you follow this, you understand why our confidence intervals for μ and our hypothesis tests about μ have been based on the magnitude of the observed value of $\overline{X}$ (for as large a sample size as possible), not on one observation!

10. An optical company uses a vacuum deposition method to apply a protective coating to certain lenses. The coating is built up one layer at a time. The thickness, X, of a given layer is a normally distributed random variable with mean $\mu = 0.5$ microns and standard deviation σ microns. Suppose that 2.28% of layers have thickness above 1.5. If a random sample of 36 layers are applied, what is the probability that their total thickness will exceed 24 microns?

A) 0.0228 **B)** 0.1 **C)** 0.1587 **D)** 0.3413 **E)** 0.5 **F)** 0.6

SOLUTION: The answer is **(A)**.
Details: We first find the value of σ. We are given that $P(X > 1.5) = 0.0228$. Standardizing as usual (see e.g. the *example (HEIGHTS)* on *page 23* of *Statistics Lecture Notes),* we have

$P(X > 0.7) = P\left(\frac{X-\mu}{\sigma} > \frac{1.5-\mu}{\sigma}\right) = P\left(Z > \frac{1.5-0.5}{\sigma}\right) = P\left(Z > \frac{1}{\sigma}\right).$ We thus have $P\left(Z > \frac{1}{\sigma}\right) = 0.0228$ and using the z-tables, we have $\frac{1}{\sigma} = 2.0,$ so $\sigma = 0.5$

The rest of the problem is similar to *example (HEIGHTS continued)* on *page 25* of *Statistics Lecture Notes* See also *Question 4* on *Problem Sheet 2* and examples in class including about an elevator breaking down!

Let X_i be the thickness of the *i*the layer, $i = 1, 2, \ldots, 36.$ The total thickness of the 36 layers is then $\sum_{i=1}^{n} X_i$ where $n = 36.$ We want $P\left(\sum_{i=1}^{n} X_i > 24)\right)$ and upon dividing each side of the inequality by 36, this is the same as asking for $P\left(\frac{\sum_{i=1}^{n} X_i}{36} > \frac{24}{36}\right),$ i.e. $P(\overline{X} > \frac{2}{3}).$ Standardizing we want

$$P(\overline{X} > \tfrac{2}{3}) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{\frac{2}{3}-\mu}{\sigma/\sqrt{n}}\right) = P\left(Z > \frac{\frac{2}{3}-\frac{1}{2}}{0.5/\sqrt{36}}\right) = P(Z > 2) \underset{\text{from z-tables}}{=} 0.0228$$