## Math 2200 Spring 2017, Exam 1

You may use any calculator. You may use ONE "cheat sheet" in the form of a 4" x 6 " note card (the medium size of the standard three sizes).

The exam is out of 100 points. The first 19 problems are worth 5 points each. The last 5 problems are worth 1 point each.

1. Consider the following incomplete frequency table and incomplete relative frequency table for the categorical variable Type of Voyager on board the Titanic on the night of 14 April 1912:

| Type of Voyager | Count |
| :--- | ---: |
| First Class | 325 |
| Second Class | 285 |
| Third Class | 706 |
| Crew |  |
| Total |  |


| Type of Voyager | Percentage |
| :--- | ---: |
| First Class |  |
| Second Class | 12.94866 |
| Third Class |  |
| Crew | 40.20900 |
| Total | 100 |

How many crew members were on the Titanic that night?
A) 880
B) 881
C) 882
D) 883
E) 884
F) 885
G) 886
H) 887
I) 888
J) 889

## Solution F) 885

The second rows of the tables are the Rosetta Stone. If $x$ represents the total number of voyagers on the Titanic that night, then $285=0.1294866 x$, or $x=2201$. The number of crew members was $2201 \times 0.40209$, or 885 .
2. This question and the next three pertain to a Swedish study that investigated a possible association between prostate cancer and fish consumption. In the study, 6272 men were followed for 30 years. Each man in the study described the contribution of fish to his overall diet by one of four levels offered by the researchers. The variable Fish Consumption Level was cross-tabulated with the variable Prostate Cancer Diagnosis, which had two values, Yes and No. The joint frequency counts can be found in the following table.

|  | No | Yes | Total |
| :--- | :---: | :---: | ---: |
| No part or nearly no part | 110 | 14 | 124 |
| A small part | 2420 | 201 | 2621 |
| A moderate part | 2769 | 209 | 2978 |
| A large part | 507 | 42 | 549 |
| Total | 5806 | 466 | 6272 |

For what percentage of the men in the study was fish not a moderate part of their diet?
A) 48.538
B) 49.865
C) 51.192
D) 52.519
E) 53.846
F) 55.173
G) 56.500
H) 57.827
I) 59.154
J) 60.481

## Solution D) 52.519

The percentage is $100(6272-2978) / 6272$, or 52.519 .
3. For each of the conditional distributions of Prostate Cancer Diagnosis arising from this cross-tabulation, find the relative frequency of men who were diagnosed with prostate cancer. Of these relative frequencies, which is largest?
A) 0.072
B) 0.113
C) 0.154
D) 0.195
E) 0.236
F) 0.277
G) 0.318
H) 0.359
I) 0.400
J) 0.441

## Solution B) 0.113

By row, the relative frequencies of men who were diagnosed with prostate cancer are: $14 / 124$, or 0.1129032 ; $201 / 2621$, or $0.07668829 ; 209 / 2978$, or $0.07018133 ; 42 / 549$, or 0.07650273 . The largest of these values is 0.1129032 .

```
In R
> crosstab.data = c(110,14, 2420, 201, 2769, 209, 507, 42)
> crosstab.matrix = matrix(crosstab.data, nrow = 4, byrow = TRUE)
> crosstab.matrix
        [,1] [,2]
[1,] 110 14
[2,] 2420 201
[3,] 2769 209
[4,] 507 42
> prop.table(crosstab.matrix, 1)
    [,1] [,2]
[1,] 0.8870968 0.11290323
[2,] 0.9233117 0.07668829
[3,] 0.9298187 0.07018133
[4,] 0.9234973 0.07650273
> max( (prop.table(crosstab.matrix, 1))[,2] )
[1] 0.1129032
```

4. The two problems on this page use the contingency table presented on the preceding page. We reproduce it here for ease of reference.

|  | No | Yes | Total |
| :--- | :---: | :---: | ---: |
| No part or nearly no part | 110 | 14 | 124 |
| A small part | 2420 | 201 | 2621 |
| A moderate part | 2769 | 209 | 2978 |
| A large part | 507 | 42 | 549 |
| Total | 5806 | 466 | 6272 |

Determine the percentage $P_{1}$ of the surveyees who had diets to which fish contributed at least a small part. Of the surveyees who were not diagnosed with prostate cancer, determine the percentage $P_{2}$ who had diets to which fish contributed at least a small part. Subtract the first percentage $P_{1}$ from the second percentage $P_{2}$ and answer with this difference.
A) -0.2577
B) -0.2010
C) -0.1443
D) -0.0876
E) -0.0309
F) 0.0258
G) 0.0825
H) 0.1392
I) 0.1959
J) 0.2526

## Solution G) 0.0825

We have $P_{1}=100(6272-124) / 6272$, or $P_{1}=98.02296$. Also, $P_{2}=100(5806-110) / 5806$, or $P_{2}=98.10541$. The requested difference is $P_{2}-P_{1}=98.10541-98.02296=0.08245$.
5. The table presented in problem 2 and reproduced above contains some evidence that fish consumption might provide some protection against the development of prostate cancer. Often, when a food item seems to have some preventive benefit, researchers are interested in whether there is a threshold level beyond which additional consumption has no further benefit. Suppose the Swedish researchers had studied only men who had diets to which fish contributed only a small or moderate part, the aim being to determine if an increase from the lower level (small part of the diet) to the higher level (moderate part of the diet) had any additional benefit in preventing prostate cancer.

A pair of unequal numbers results in a positive difference when the smaller number is subtracted from the larger number. Using the two conditional distributions of Prostate Cancer Diagnosis, the researchers would
have looked for two pairs of unequal numbers as an indicator of a dependence between prostate cancer and the two levels of fish consumption. What positive difference would they have found?
A) 0.0013
B) 0.0026
C) 0.0039
D) 0.0052
E) 0.0065
F) 0.0078
G) 0.0091
H) 0.0104
I) 0.0117
J) 0.0130

## Solution E) 0.0065

The middle two rows of the contingency table expressed as relative frquencies give the following table of the conditional distributions of Prostate Cancer Diagnosis for each of the two levels of fish consumption that we are interested in:

|  | No | Yes | Total |
| :--- | :---: | :---: | ---: |
| A small part | 0.9233 | 0.0767 | 1.0 |
| A moderate part | 0.9298 | 0.0702 | 1.0 |

The absolute difference of the two entries in both columns is 0.0065 .
As a continuation of the R session shown in the solution to Problem 3, we have

```
> abs( (prop.table(crosstab.matrix, 1))[2,] - (prop.table(crosstab.matrix, 1))[3,] )
[1] 0.006506957 0.006506957
```

6. On a midterm exam with 100 questions, Barunka guessed on 60 questions and exactly 20 of her guesses were correct. On the same exam, Baruska guessed on 11 questions and exactly 4 of her guesses were correct. On the final exam with 100 questions, Barunka guessed on 5 questions and exactly 1 of her guesses was correct. On the same exam, Baruska guessed on 54 questions and exactly $x$ of her guesses were correct. Suppose that the overall success rates of Barunka and Baruska display Simpson's Paradox. Of the 10 values listed below, which is the largest that $x$ might have been.
A) 8
B) 10
C) 12
D) 14
E) 16
F) 18
G) 20
H) 22
I) 24
J) 26

## Solution E) 16

There are two aspects of Simpson's Paradox: 1) Each component trends in the same direction, and 2) The aggregate is opposite to the trend of the components. Each aspect puts a bound on $x$. Barunka's guess rate of $20 / 20$, or $1 / 3$, on the first exam is less than Baruska's guess rate of $4 / 11$. For there to be an example of Simpson's Paradox, this trend must be repeated on the second exam. Therefore, $1 / 5<x / 54$. We deduce that $x \geq 11$. For the second aspect of Simpson's Paradox, the reversal of the trend in the aggregate, we must have $(20+1) /(60+5)>(4+x) /(11+54)$, or $21 / 65>(4+x) / 65$, or $21>x+4$, or $x<17$. Thus, $x$ might have the value $11,12,13,14,15$, or 16 . The largest of these is in the answer list: 16 .
7. Every even-numbered year, each surveyee in the General Social Survey is asked his or her opinion concerning societal assistance to the poor. The incomplete table below records the counts of responses for the three indicated years.

|  | 2010 | 2012 | 2014 | Total |
| :--- | ---: | ---: | ---: | ---: |
| Too little |  | 587 |  | 2050 |
| About right | 223 |  | 301 |  |
| Too much | 97 | 100 |  |  |
| Don't know | 34 |  | 31 | 106 |
| Total |  | 976 | 1264 | 3270 |

In the year 2014, how many respondents thought society was doing too much?
A) 137
B) 141
C) 145
D) 149
E) 153
F) 157
G) 161
H) 165
I) 169
J) 173

## Solution C) 145

The missing bottom margin entry is $3270-(976+1264)$, or 1030 . The missing tabular entry in the first column is $1030-(223+97+34)$, or 676 . The missing entry in the second row is $2050-(676+587)$, or 787 . The missing entry in the 3 'rd row, 3 'rd column is $1264-(787+301+31)$, or 145 , which is the requested answer. The complete table is

|  | 2010 | 2012 | 2014 | Total |
| :--- | ---: | ---: | ---: | ---: |
| Too little | 676 | 587 | 787 | 2050 |
| About right | 223 | 248 | 301 | 772 |
| Too much | 97 | 100 | 145 | 342 |
| Don't know | 34 | 41 | 31 | 106 |
| Total | 1030 | 976 | 1264 | 3270 |

8. Every time that a loonie is tossed and lands loon side up, a 1 is recorded. When the coin lands Queen side up, a -1 is recorded. A loonie never lands on its edge. After 24 tosses, the number of positive values in the distribution equals the number of negative values. What is the standard deviation of the distribution?
A) 0.9914 ,
B) 0.9957
C) 1.0000
D) 1.0043
E) 1.0086
F) 1.0129
G) 1.0172
H) 1.0215
I) 1.0258
J) 1.0301

## Solution H) 1.0215

The sum of all the observations is the sum of 24 ones and 25 minus ones. Therefore, the mean $\bar{x}$ is $0 / 24$, or 0 . The square of each deviation is $( \pm 1-0)^{2}$, or 1 . The sum of the squared deviations over all observations is 24 Therefore, the standard deviation of the distribution is $\sqrt{24 /(24-1)}$, or 1.021508 .
9. To obtain the lower quartile of the distribution $X=a, b, c, d$, $e$, which has been given in sorted order (low to high), Miles Prower used the course definition and correctly obtained the value 13. Wile E used an Acme brand calculator, which defines the lower quartile of X to be the median of the numbers that precede the median of X (in the given low-to-high sorting). Wile E's Acme calculator displayed 12. What is the value of the first number in the five number summary of X ?
A) 8.5
B) 9
C) 9.5
D) 10
E) 10.5
F) 11
G) 11.5
H) 12
I) 12.5
J) 13

## Solution F) 11

The median of X is c . It is a middle value of X , so it is included with a and b in calculating the lower quartile (according to course conventions). The median of $\mathrm{a}, \mathrm{b}, \mathrm{c}$ is b . Therefore, $\mathrm{b}=13$. The Acme brand calculator discards $c$ in its calculation of the lower quartile and displays $(a+b) / 2$ for the lower quartile. Therefore, $(\mathrm{a}+\mathrm{b}) / 2=12$, or $\mathrm{a}+\mathrm{b}=24$, or $\mathrm{a}+13=24$, or $\mathrm{a}=11$. Because a is the smallest value in X , it is the first number of the five number summary of X .
10. This problem and the next four pertain to the following sorted observations, which represent the number of assists Wayne Gretzky, aka The Great One, obtained in each of his 20 NHL seasons:

374953677279869092102109109114118120121122125135163
To draw a histogram, this data is binned so that

1) The class width is an integer.
2) The least lower class limit is 37 .
3) The observed value 125 is a class limit.
4) The rightmost bin is nonempty.
5) The number of bins is more than 3 and less than 16 .

What is the largest upper class limit?
A) 163
B) 164
C) 165
D) 166
E) 167
F) 168
G) 169
H) 170
I) 171
J) 172

Solution
G) 169

As we will see, there are two allowable binnings that meet the requirements. Nevertheless, the answer to this problem is unique: both allowable binnings have the same greatest class limit.

Let $h$ be the class width. By condition (1), $h$ is an integer. Conditions (2) and (3) tell us that there is an integer $k$ such that $125=37+h \cdot k$. Thus, $h \cdot k=88$. Thus, $h$ is a divisor of 88 . It follows that $h=$ $1,2,4,8,11,22,44$, or 88 . Clearly $h=1,2$, or 4 result in more than 15 bins. Also, $37+8 \cdot 15=157<163$, so $h=8$ also requires more than 15 bins. It follows that $h=11,22,44$, or 88 . However, if $h=44$, then $37+44 \cdot 3=169$, which shows that there are 3 bins. So we rule out $h=44$. It follows that 88 is also out, and, in any event, we see that there are only 3 bins for $h=88$. We are left with $h=11$ or $h=22$.

Suppose that $h=11$. If $k$ is the number of bins, then we require $37+11 k>163$, or $k>126 / 11$. The smallest integer that satisfies this equation is $k=12$, which is more than 2 and less than 16 . The class limits are $37,48,59,70,81,92,103,114,125,136,147,158,169$.

Suppose that $h=22$. If $k$ is the number of bins, then we require $37+22 k>163$, or $k>126 / 22$. The smallest integer that satisfies this equation is $k=6$, which is more than 2 and less than 16 . The class limits are $37,59,81,103,125,147,169$.

Either way, the largest class limit is 169 .
11. When Gretzky's assists per year are binned according to the specifics stated in the preceding problem (and according to the general conventions we have adopted for the course), the resulting histogram is
A) uniform
B) multimodal with more than three modes
C) multimodal with three modes
D) bimodal with equal height modes
E) bimodal with greater left mode
F) bimodal with greater right mode
G) unimodal symmetric
H) unimodal skewed left
I) unimodal skewed right
J) not described by any of the preceding shapes

## Solution C) multimodal with three modes OR H) unimodal skewed left

Here is the R code for the two allowable histograms:

```
> Gretzky.assists = c(37,49,53,67,72,79,86,90,92,102,109,109,114,118,120,121,122,125,135,163)
> hist(Gretzky.assists, breaks = 37 + 11*(0:12), right=FALSE) # bin width = 11, 12 bins
> hist(Gretzky.assists, breaks = 37 + 22*(0:6), right=FALSE) # bin width = 22, 6 bins
```

First, let's look at the 12 bin histogram with class width 11:


Histogram of Wayne Gretzky's Assists per Season, Bin Width: 11, Number of Bins: 12
Whether this histogram shows a distribution that is unimodal and skewed left, or a multimodal distribution with three modes is debatable. The book-end modes look more dramatic than they really are: each of these two modes exceeds its bin neighbors by only one observation. To avoid debate, both (C) and (H) are deemed correct. The 6 bin histogram with class width 22 is less controversial: it is unimodal and skewed left.


Histogram of Wayne Gretzky's Assists per Season, Bin Width: 22, Number of Bins: 6
12. The problems on this page continue with the distribution of Gretzky's assists per year, which is repeated here for ease of reference:

374953677279869092102109109114118120121122125135163
The IQR of the distribution of Gretzky's assists per year is
A) 43
B) 43.5
C) 44
D) 44.5
E) 45
F) 45.5
G) 46
H) 46.5
I) 47
J) 47.5

## Solution E) 45

The five number summary is $37,75.5,105.5,120.5,163$. The IQR is $(120.5-75.5)$, or 45 .
13. For a horizontal Tukey boxplot for the distribution of Gretzky's assists per year, the fence on the right is at what value?
A) 170
B) 172
C) 174
D) 176
E) 178
F) 180
G) 182
H) 184
I) 186
J) 188

## Solution J) 188

The fence on the right is at $120.5+1.5 \times \mathrm{IQR}$, or $120.5+(1.5)(45)$, or 188 .
For the next question, the fence on the left is at $75.5-1.5 \times \mathrm{IQR}$, or $75.5-(1.5)(45)$, or 8 .
14. The length of the whisker on the right side of a horizontal Tukey boxplot for the distribution of Gretzky's assists per year exceeds the length of the left whisker by what amount?
A) 1
B) 2
C) 3
D) 4
E) 5
F) 6
G) 7
H) 8
I) 9
J) 10

## $\begin{array}{ll}\text { Solution } & \text { D) } 4\end{array}$

There are no observations outside the fences. Therefore, $163-120.5$, or 42.5 is the length of the right whisker, and $75.5-37$, or 38.5 is the length of the left whisker. The right exceeds the left by 4 .
15. The cholesterol levels of a population of American women can be described by a normal model with mean $188 \mathrm{mg} / \mathrm{dL}$ and standard deviation of $24 \mathrm{mg} / \mathrm{dL}$. The women with cholesterol levels between 150 and 170 $\mathrm{mg} / \mathrm{dL}$ make up what percentage of the population? Round to the nearest integer.
A) 14
B) 15
C) 16
D) 17
E) 18
F) 19
G) 20
H) 21
I) 22
J) 23

## Solution D) 17

A cholesterol level of 170 corresponds to a z-score of $(170-188) / 24$, or -0.75 . A cholesterol level of 150 corresponds to a z-score of $(150-188) / 24$, or -1.5833 . The required percentage is

$$
\begin{aligned}
100(\Phi(-0.75)-\Phi(-1.5833)) & =100((1-\Phi(0.75))-(1-\Phi(1.5833))) \\
& =100(\Phi(1.5833)-\Phi(0.75)) \\
& =100(0.9433-0.7734) \\
& =16.99 .
\end{aligned}
$$

16. Refer to the model in the preceding problem. Above what value in $\mathrm{mg} / \mathrm{dL}$ are the highest $15 \%$ of women's cholesterol levels?
A) 206.12
B) 206.87
C) 207.62
D) 208.37
E) 209.12
F) 209.87
G) 210.62
H) 211.37
I) 212.12
J) 212.87

## Solution J) 212.87

The z-score coresponding to 0.85 is 1.0364 . The cholesterol level $x$ that has z -score equal to 1.0364 satisfies $(x-188) / 24=1.0364$, or $x=188+(24)(1.0364)$, or $x=212.8736$.
17. The time in months from birth until a baby begins to walk can be described by a normal model. By the age of 10 months, only $5 \%$ of babies have learned to walk. By the age of 13 months, only $25 \%$ of babies
have not started walking. This problem and the next concern the unspecified mean and standard deviation of the normal model. To answer this problem, calculate and state the standard deviation.
A) 0.7874
B) 0.8597
C) 0.9320
D) 1.0043
E) 1.0766
F) 1.1489
G) 1.2212
H) 1.2935
I) 1.3658
J) 1.4381

## Solution H) 1.2935

Let $\mu$ be the unknown mean and let $\sigma$ be the unknown standard deviation. Let X be the time from birth to learning to walk. Then $Z=(X-\mu) / \sigma$ is standard normal. We have $\Phi(z)=0.05$ for the value $z$ such that $\Phi(-z)=0.95$. Thus, $-z=1.644854$, or $z=-1.644854$. It follows that $(10-\mu) / \sigma=-1.644854$, or $10-\mu=-1.644854 \sigma$. Similarly, $\Phi(z)=0.75$ for $z=0.6744898$ tells us that $(13-\mu) / \sigma=0.6744898$, or $13-\mu=0.6744898 \sigma$. Thus, $(13-\mu)-(10-\mu)=3$ and $(13-\mu)-(10-\mu)=0.6744898 \sigma-(-1.644854 \sigma)=$ $2.319344 \sigma$. As a result, $3=2.319344 \sigma$, or $\sigma=3 / 2.319344$, or $\sigma=1.2935$.
18. To answer this problem, calculate and state the mean of the normal model described in the preceding problem.
A) 11.5812
B) 11.6495
C) 11.7178
D) 11.7861
E) 11.8544
F) 11.9227
G) 11.9910
H) 12.0593
I) 12.1276
J) 12.1959

## Solution I) 12.1276

Using the value $\sigma=1.2935$ from the last problem, and either the equation $(10-\mu) / \sigma=-1.644854$ or the equation $13-\mu=0.6744898 \sigma$ from the last problem, we obtain $m=12.1276$.
19. All volume measurements in this problem are in cubic centimeters. Assume that the volume of an adult male mouth can be described by a normal model. Assume that the volume of an adult female mouth can also be described by a normal model. A Cornell researcher who measured mouth volumes determined that the mean and standard deviation for the male population are 66 and 17 respectively, and the mean and standard deviation for the female population are 54 and 14.5 respectively. Among the men whose measurements she recorded, the largest mouth volume was 111.2. That was one big mouth. How large would a comparable female mouth be? (Male and female mouths are comparable if the percentages of smaller mouths in their respective populations are equal.)
A) 92.55
B) 93.48
C) 94.41
D) 95.34
E) 96.27
F) 97.20
G) 98.13
H) 99.06
I) 99.99
J) 100.92

## Solution A) 92.55

The z-score of the male big mouth was $(111.2-66) / 17$, or 2.6588 . The volume measurement $y$ of a female mouth with the same z-score satisfies $(y-54) / 14.5=2.6588$, or $y=54+(14.5)(2.6588)$, or $y=92.5526$.
20. Which of the following numbers is the return of

$$
(1.234+0.123 *(-2: 10))[5]
$$

in R ?
A) 1.357
B) 1.480
C) 1.603
D) 1.726
E) 1.849
F) 1.972
G) 2.095
H) 2.218
I) 2.341
J) 2.464

## Solution B) 1.480

In $R,-2: 10$ is the vector $-2,-1,0,1,2, \ldots, 10$. Therefore $0.123 *(-2: 10)$ is the vector $-0.246,-0.123,0$, $0.123,0.246, \ldots, 1.230$, and $1.234+0.123 *(-2: 10)$ is the vector $0.988,1.111,1.234,1.357,1.480, \ldots$, 2.464. The code $(1.234+0.123 *(-2: 10))[5]$ picks out the fifth entry of this list, namely 1.480 .

```
> (1.234 + 0.123*(-2:10))[5]
```

[1] 1.48
21. What is the return in R of the following code?
> $\mathrm{M}=$ matrix ( $1: 12$, ncol $=3$, byrow = TRUE)
$>\operatorname{sum}(M[3]$,
A) 16
B) 17
C) 18
D) 19
E) 20
F) 21
G) 22
H) 23
I) 24
J) 25

## Solution I) 24

The input $M=\operatorname{matrix}(1: 12, \mathrm{ncol}=3$, byrow $=$ TRUE) creates a matrix named M with 4 rows and 3 columns. The cells are filled with the numbers $1,2,3, \ldots, 12$ by row, so that $1,2,3$ is the first row, 4,5 , 6 the second, $7,8,9$ the third. That's all we need: the call sum (M[3,]) asks for the sum of the entries in the third row, namely $7+8+9$, or 24 . Now you know why 6 was afraid of 7 .

```
> M = matrix( 1:12, ncol = 3, byrow = TRUE)
> M
[,1] [,2] [,3]
[1,] 1 2 3
[2,] 4 5 6
[3,] 7 8 9
[4,] 10 11 12
> sum(M[3,])
[1] 24
```

22. For the vector
$>X=c(1,2,2,3,3,3,4,4,4,4,5,5,5,5,5)$
in $R$, which of the following commands produced the histogram that is shown?

Histogram of $X$

A) $\operatorname{hist}(\mathrm{X}$, breaks $=0: 6$, right $=$ FALSE $)$
B) $\operatorname{hist}(\mathrm{X}$, breaks $=0: 6$, right $=$ TRUE $)$
C) $\operatorname{hist}(\mathrm{X}$, breaks $=1: 6$, right $=$ FALSE $)$
D) $\operatorname{hist}(\mathrm{X}$, breaks $=1: 6$, right $=$ TRUE $)$
E) $\operatorname{hist}(\mathrm{X}$, nodes $=0: 6$, right $=$ FALSE $)$
F) $\operatorname{hist}(\mathrm{X}$, nodes $=0: 6$, right $=$ TRUE $)$
G) $\operatorname{hist}(\mathrm{X}$, nodes $=1: 6$, right $=$ FALSE $)$
H) $\operatorname{hist}(\mathrm{X}$, nodes $=1: 6$, right $=$ TRUE $)$
I) Wrong Answer
J) Another Wrong Answer

Solution A) hist $(\mathbf{X}$, breaks $=0: 6$, right $=$ FALSE $)$

Eliminate all but the first four answers. The term for "class limits" used in other area of mathematics (such as Numerical Methods) can be "nodes", but in R's histogram call, the term for bin endpoints is "breaks". From the plot we see that the binning begins at 0 . We have narrowed the answer down to A or B . Of these, A will place every observation that is a class limit into the bin on the right. (This can seem counterintuitive, but the argument right $=$ FALSE means that it is false for a bin to be right-closed). That means the bin $[1,2)$ has the observation 1 , the bin $[2,3)$ has the observation 2,2 , the bin $[3,4)$ has the observation $3,3,3$ and so on. The R input hist ( X , breaks $=0: 6$, right $=$ FALSE) given in answer A is therefore consistent with the given histogram. By contrast, the call hist (X, breaks $=0: 6$, right $=$ TRUE) in answer B has bins that are left-open and right-closed. The bar above ( 0,1 ] has height 1 , and the bar above $(5,6]$ has height 0 .
23. Which of the following numbers is the return for

```
> pnorm(12, mean = 7, sd = 4, lower.tail = TRUE)
```

in R ?
A) 0.8708
B) 0.8749
C) 0.8790
D) 0.8830
E) 0.8869
F) 0.8907
G) 0.8944
H) 0.8980
I) 0.9015
J) 0.9049

## Solution G) 0.8944

In a normal distribution with mean 7 and standard deviation 4 , the $z$-score of 12 is $(12-7) / 4$, or 1.25 . The call pnorm(12, mean $=7$, sd $=4$, lower.tail $=$ TRUE) therefore asks for $\Phi(1.25)$, which is the entry 0.8944 in the $\Phi$ table in the row for 1.2 and the column for 0.05 .
24. Which of the following numbers is the return for

```
> qnorm(0.0516, lower.tail = FALSE)
```

in R ?
A) 1.59
B) 1.60
C) 1.61
D) 1.62
E) 1.63
F) 1.64
G) 1.65
H) 1.66
I) 1.67
J) 1.68

## Solution E) 1.63

The call qnorm ( 0.0516 , lower.tail $=$ FALSE $)$ specifies the value of z such that the area under the bell curve to the right of z is 0.0516 . In other words, the call is for the value of z such that the area under the bell curve to the left of z is $1-0.0516$, or 0.9484 . That is R's return is the value of z for which $\Phi(z)=0.9484$. This number is in in the $\Phi$ table - the row for 1.6 and the column for 0.03 . Tha answer is $1.6+0.03$, or 1.63. Here is a copy-and-paste from an R session:

```
> qnorm(0.0516, lower.tail = FALSE)
```

[1] 1.629534

## STANDARD STATISTICAL TABLES

## 1. Areas under the Normal Distribution

The table gives the curulative probability up to the standardised nornal value 2
$P[Z<Z]=\int_{-\infty}^{2} \frac{1}{\sqrt{2 \pi}} \exp \left(-\frac{1}{2} Z^{2}\right) d Z$


| 2 | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7020 | 0.7054 | 0.7089 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.84 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.855 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8804 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 0 | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| , | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9865 | 0.9868 | 0.9871 | 0.9874 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9924 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9980 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| 2 | 3.00 | 3.10 | 3.20 | 3.30 | 3.40 | 3.50 | 3.60 | 3.70 | 3.80 | 3.90 |
| P | 0.9986 | 0.9990 | 0.9993 | 0.9995 | 0.9997 | 0.9998 | . 99 | 0.9999 | . 9999 | . 0000 |

