Write your full name legibly at the top & the first letter of your last name in the box (1 pt).
Put your calculator's number on the blank below the box.
Written answers should be concise and final answers should be copied into blanks provided where designated.

MULTIPLE CHOICE (3 pts each)  
Exam is worth 150 points

(1) Consider a population of black and white discs that has a certain proportion of black discs we designate "J". What is the relationship between J and the variance of estimates of J as determined from samples.
(A) As J increases, the variance of the estimates always decreases.
(B) As J increases, the variance of the estimates always increases.
(C) The closer to 0.5 that J is, the higher the variance of the estimates.
(D) The closer to 0.5 that J is, the smaller the variance of the estimates.
(E) The variance of the estimates depends on the size of the samples, not the value of J.

(2) If a population exhibits a skew of 0.38 and a kurtosis of 3.54 which of the following is true.
(A) The distribution of population values is left skewed and platykurtic.
(B) The distribution of population values is left skewed and leptokurtic.
(C) The distribution of population values is right skewed and platykurtic.
(D) The distribution of population values is right skewed and leptokurtic.
(E) The distribution of population values is binomial.

(3) If a population exhibits a skew of 0.38 and a kurtosis of 3.54 which of the following is true.
(A) The population mean exceeds the median.
(B) The population median exceeds the mean.
(C) The population standard deviation exceeds the variance.
(D) The population variance exceeds the standard deviation.
(E) The population median equals the mean.

(4) Which of the following is true for a data set exhibiting a Poisson distribution?
(A) Mean and variance are equal.  
(D) Distribution exhibits independence.
(B) Median and mean are equal.  
(E) Distribution is symmetric.
(C) Median and variance are equal.

(5) Which of the following is NOT a good description of at least part of our approach to descriptive statistics:
(A) We seek to define values that represent properties of the data distribution.
(B) We seek to determine the location, spread and shape of data distributions with summary statistics.
(C) We seek to determine whether the sample statistics match the population parameters.
(D) We seek to condense large amounts of data into smaller sets of values.
(E) We seek to work with graphical tools to present numerical data in a visual manner.

(6) Which of the following is NOT an assumption required in order for a distribution of observations to have a binomial distribution?
(A) Each trial is independent of all others.
(B) Each trial is mutually exclusive of all others.
(C) Results of trials can be defined to have only two possibilities.
(D) The probability of observing results doesn't change over time.
(E) There are a known number of trials.
(7) Which of the following sets of data most likely exhibits a Poisson distribution?
(A) The number of victories the LA Lakers have each year.
(B) The number of losses the LA Clippers have each year.
(C) The mean temperature of California on June 5th each year.
(D) The number of meteors that land in California each year.
(E) The number of deaths of California citizens for each age in any given year.

(8) The proportion of observations in a data set that are below the median is:
(A) 50%    (B) 66%    (C) 75%    (D) 95%    (E) 99%

(9) In general, the presence or absence of outliers has the most effect on the consistency of which of the following statistics?
(A) IQR    (B) Mode    (C) Median    (D) Q1    (E) Variance

(10) Which of the following is NOT a measurement of location?
(A) Kurtosis    (B) Mean    (C) Median    (D) Mid-range    (E) Mode

(11) Which is the following an accurate description of our definitions of “independent” and “mutually exclusive”?
(A) A and B are independent if p(A) = p(B) and mutually exclusive if p(A or B) = 0.
(B) A and B are independent if p(A) = p(B) and mutually exclusive if p(A and B) = 0.
(C) A and B are independent if p(A) = p(A|B) and mutually exclusive if p(A and B) = 0.
(D) A and B are independent if p(A) = p(B) and mutually exclusive if p(A or B) = 0.
(E) A and B are independent if p(A) = p(A|B) and mutually exclusive if p(A or B) = 1.

(12) Which of the following is NOT correct?
(A) The normal distribution approximates the binomial distribution when the number of trials is large and the probability of each success is 0.5.
(B) The Gaussian distribution approximates the binomial distribution when the number of trials is large and the probability of each success is 0.5.
(C) The Gaussian distribution is equivalent to the normal distribution.
(D) The Poisson distribution approximates the binomial distribution when the number of trials is large and the probability of each success is 0.5.
(E) The Poisson distribution approximates the binomial distribution when the number of trials is large and the probability of each success is small.

(13) The symbolism A~N(5,12) is shorthand for which of the following?
(A) The data set "A" has a normal distribution with a mean value of 5 and a standard deviation of 12.
(B) The data set "A" has a normal distribution with a mean value of 5 and a variance of 12.
(C) The data set "A" has a normal distribution with a mean value of 12 and a standard deviation of 5.
(D) The data set "A" has a normal distribution with a mean value of 12 and a standard deviation of 5.
(E) The data set "A" has values that range from a minimum of 5 to a maximum of 12.

(14) Which is the best way of thinking about when two events A and B are independent?
(A) The probability of A occurring is the same as the probability of B occurring.
(B) The probability of A occurring is complementary to the probability of B occurring.
(C) The probability of A occurring is not complementary to the probability of B occurring.
(D) Whether A does or does not occur influences the probability of B occurring.
(E) Whether A does or does not occur does not influence the probability of B occurring.
The next 3 questions are based on the list of cognitive biases that you were instructed to read for this exam.

(15) The website described this bias as the "if it ain't broke, don't fix it" bias and it leads to people not acting even though there are many reasons to act.
(A) Confirmation Bias  (C) Ingroup Bias  (E) Status-Quo Bias
(B) Current Moment Bias  (D) Negativity Bias

(16) The world is a dangerous place ... or is it? Within the last month there was a violent shooting near UCSB and Boko Haram kidnapped a few hundred female students from their school. Over the past hundred years however the US rates of violent murder and global deaths due to warfare have actually been decreasing and are lower now than at virtually any point in history. The bias that causes us to believe that the world is more dangerous when it is actually safer is termed which of the following?
(A) Anchoring Effect  (D) Observational Selection Bias
(B) Bandwagon effect  (E) Status-Quo Bias
(C) Negativity Bias

(17) World of Warcraft players often "run instances", fighting through the same part of the game again and again to defeat a particular opponent because they have a chance of rewarding the player with a specific item they desire. The chance of getting each specific item from an opponent is a constant percentage after each battle. Many players run instances over and over again telling themselves each time that since the previous run was unsuccessful then next one must be more likely to be successful and they will get the item they desire. Which of the following biases contributes to the long late nights spent running the same instance over and over for this reason?
(A) Confirmation Bias  (D) Neglecting probability
(B) Gambler's fallacy  (E) Post-purchase rationalization
(C) Ingroup bias

These 3 questions are inspired by the XKCD.com cartoon published just last week on May 28, 2014 and shown to the right. Imagine that you're drawing at random from an urn containing fifteen balls - six red and nine black.

(18) If you pull out five balls and three are red while two are black how many different ways could this have occurred?
(A) 2  (C) 6  (E) 10
(B) 4  (D) 8

(19) If you pull out four balls without replacement, which of the following is closest to the probability of all the balls being black?
(A) 0.06  (C) 0.12  (E) 0.18
(B) 0.09  (D) 0.15

(20) If you pull out two balls without replacement, which of the following is closest to the probability of the balls being different colors?
(A) 0.2  (C) 0.4  (E) 0.6
(B) 0.3  (D) 0.5
(21) POISSON DISTRIBUTION (3 pts each). Consider a situation in which a researcher has set up an automated camera that records endangered owls as they pass the camera at night (they are nocturnal so filming them is otherwise difficult). Long term data indicates that the mean number of owls that pass by the camera each night is 4. Use a value of $e=2.7$ for your calculations and provide all answers to the nearest 0.001.

(a) What is the probability that the camera does not record an owl during a given night? 
prob = __________

(b) What is the probability that the camera records 3 owls during a given night? 
prob = __________

(b) What is the probability that the camera records more than 4 owls during a given night? 
prob = __________

(22) CALCULATIONS (3 pts each). Consider a data set of 1000 values that exhibits a normal distribution in which the mean is 100 and the standard deviation is 10. For the following questions, provide all non-integer values to nearest 0.01.

(a) Approximately what proportion of the data is less than 90? 
prop = ________ %

(b) If a random value is chosen from this data set, what is the probability that it will be smaller than 103.3?
prob = ________ %

(c) How many values in the data set are between 94 and 102? 
# values = ________

(d) What is the value of Q1? 
Q1 = ________

(e) What is the value of the data set's IQR? 
IQR = ________

(f) What is the Coefficient of variation of the data set? 
CV = ________

(h) What is the skew of the data set? 
Skew = ________

(g) What is the mid-range of the data set? 
Mid-range = ________
(23) **PROBABILITY** (4 pts each). Consider a lake full of three species of fish: walleye, bass and perch. The lake contains 800 walleye, 800 bass and 400 perch. (This population is large enough that you can consider all samples to be sampling with replacement for any calculations). Provide all non-integer values to nearest 0.0001.

(A) What is the probability of catching a random fish and it is a bass? \[ p(B) = \ldots \]

(B) What is the probability that you catch four random fish and all four of them are perch? \[ p(4P) = \ldots \]

(C) What is the probability that you catch two random fish and they are both from the same species? \[ p(2 \text{ same}) = \ldots \]

(D) What is the probability that if you catch two fish, you catch a perch and a walleye? \[ p(P\&W) = \ldots \]

(E) What is the probability that if you catch two fish, first you catch a perch and then a walleye? \[ p(P, W) = \ldots \]

(F) What is the probability that if you catch 8 random fish you will get 5 walleye and 3 non-walleye fish? \[ p(5W, 3\text{ non}) = \ldots \]

(G) Now imagine that a disease strikes and kills all the bass. When you fish in this newly depauperate lake, what is the probability of catching a random fish and it is a bass? \[ p(B) = \ldots \]
(24) **CALCULATING STATISTICS.**

For the questions below use the sample of data shown to the right.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

(a, 2 pts each) Using all the values from the data above calculate the indicated values and fill in the blanks below. (provide all non-integer values to nearest 0.001)

Mean =  
Variance =  
Median =  
Standard deviation =  
Mode =  
Coefficient of variation =  %  
Mid-range =  
Interquartile range (IQR) =  

Describe the symmetry of this data set; the data set is  

Describe the "peakness" of this data set; the data set is  

(b, 1 pt each) If this data were to be treated as a population instead, indicate how each of the values would change (circle one option for each - inc = increase, dec = decrease, same = stays the same)

Mean: inc dec same  
Variance: inc dec same  
Median: inc dec same  
Standard deviation: inc dec same  
Mode: inc dec same  
Coefficient of variation: inc dec same  
Mid-range: inc dec same  
Interquartile range (IQR): inc dec same  

### TABLE OF Z SCORES

Table of probabilities of the standard normal distribution. Table shows the probability that a standard normal variate will have a value less than or equal to z.

<table>
<thead>
<tr>
<th>z</th>
<th>0.0013</th>
<th>0.0013</th>
<th>0.0013</th>
<th>0.0012</th>
<th>0.0012</th>
<th>0.0011</th>
<th>0.0011</th>
<th>0.0011</th>
<th>0.0010</th>
<th>0.0010</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3.0</td>
<td>0.9913</td>
<td>0.9913</td>
<td>0.9913</td>
<td>0.9912</td>
<td>0.9912</td>
<td>0.9911</td>
<td>0.9911</td>
<td>0.9911</td>
<td>0.9910</td>
<td>0.9910</td>
</tr>
<tr>
<td>-2.9</td>
<td>0.9919</td>
<td>0.9918</td>
<td>0.9918</td>
<td>0.9917</td>
<td>0.9917</td>
<td>0.9916</td>
<td>0.9916</td>
<td>0.9916</td>
<td>0.9915</td>
<td>0.9915</td>
</tr>
<tr>
<td>-2.8</td>
<td>0.9926</td>
<td>0.9925</td>
<td>0.9925</td>
<td>0.9924</td>
<td>0.9924</td>
<td>0.9923</td>
<td>0.9923</td>
<td>0.9923</td>
<td>0.9922</td>
<td>0.9922</td>
</tr>
<tr>
<td>-2.7</td>
<td>0.9935</td>
<td>0.9934</td>
<td>0.9934</td>
<td>0.9933</td>
<td>0.9933</td>
<td>0.9932</td>
<td>0.9932</td>
<td>0.9932</td>
<td>0.9931</td>
<td>0.9931</td>
</tr>
<tr>
<td>-2.6</td>
<td>0.9947</td>
<td>0.9945</td>
<td>0.9945</td>
<td>0.9944</td>
<td>0.9944</td>
<td>0.9943</td>
<td>0.9943</td>
<td>0.9943</td>
<td>0.9942</td>
<td>0.9942</td>
</tr>
<tr>
<td>-2.5</td>
<td>0.9962</td>
<td>0.9960</td>
<td>0.9960</td>
<td>0.9959</td>
<td>0.9959</td>
<td>0.9958</td>
<td>0.9958</td>
<td>0.9958</td>
<td>0.9957</td>
<td>0.9957</td>
</tr>
<tr>
<td>-2.4</td>
<td>0.9978</td>
<td>0.9976</td>
<td>0.9976</td>
<td>0.9975</td>
<td>0.9975</td>
<td>0.9974</td>
<td>0.9974</td>
<td>0.9974</td>
<td>0.9973</td>
<td>0.9973</td>
</tr>
<tr>
<td>-2.3</td>
<td>0.9992</td>
<td>0.9990</td>
<td>0.9990</td>
<td>0.9989</td>
<td>0.9989</td>
<td>0.9988</td>
<td>0.9988</td>
<td>0.9988</td>
<td>0.9987</td>
<td>0.9987</td>
</tr>
<tr>
<td>-2.2</td>
<td>0.9998</td>
<td>0.9996</td>
<td>0.9996</td>
<td>0.9995</td>
<td>0.9995</td>
<td>0.9994</td>
<td>0.9994</td>
<td>0.9994</td>
<td>0.9993</td>
<td>0.9993</td>
</tr>
<tr>
<td>-2.1</td>
<td>0.9999</td>
<td>0.9998</td>
<td>0.9998</td>
<td>0.9997</td>
<td>0.9997</td>
<td>0.9996</td>
<td>0.9996</td>
<td>0.9996</td>
<td>0.9995</td>
<td>0.9995</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9998</td>
<td>0.9998</td>
<td>0.9997</td>
<td>0.9997</td>
<td>0.9997</td>
<td>0.9996</td>
<td>0.9996</td>
</tr>
</tbody>
</table>

**Diagram:** Normal distribution curve with z values and corresponding probabilities.