Name (Last, First): ______________________________ ID: ____________ Quiz section or time: _____

10 + 15.5

Stat/Math 390, Winter, Test 1, Jan. 27, 2017; Marzban

CLOSED everything. Check front and back page. ONLY a half-size "cheat sheet" is allowed.

FIRST PAGE: give answers on these pages. DO NOT EXPLAIN. There is wrong-answer penalty.

THE REST: SHOW answer & WORK on these pages; NO CREDIT FOR CORRECT ANSWER WITHOUT EXPLANATIO

Points

1 lect3 **1.** Which of the following situations suggests that x has a truly exponential distribution:
a) The hist of x is linear when the frequency is on a log-scale. → $\log f = \alpha + \beta x \Rightarrow f = e^{\alpha+\beta x}$ = expon.
b) The frequency hist of $log(x)$ is linear.
c) The hist of $log(x)$ is linear when the frequency is on log-scale.

1.5 lect4,10 **2.** For which of the following random variables one canNOT compute a distribution mean? −0.5 penalty
a) Letter grades (A,B,...) in a class   qualitative   c) The state of a die (1,2,...6)
b) The state (H/T) of a fair coin   d) Political party affiliation of people (Demo/Repub/Indep)

1 1.62 Spr 12 **3.** Let x denote the distance (m) that an animal moves from its birth site to the first territorial vacancy it encounters. Suppose that for rats, x has an exponential distribution with parameter $\lambda = 0.01386$. Then $\mu_x = 1/\lambda$ and it represents
a) the average number of rats who move distance x.   c) the average distance moved by rats.
b) the average number of rats who move.   d) None of the above.

1.5 lect5 **4.** Which of the following distributions can be bell-shaped (skewed or not)? −0.5 penalty
a) Exponential   b) Poisson   c) Binomial   d) Normal

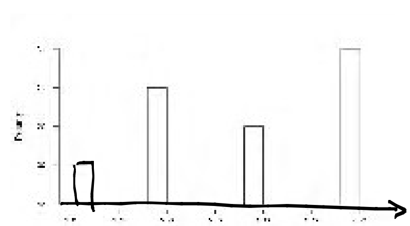
1 class question **5.** For a binomial random variable with parameters n and π, the number of levels is
a) 2   b) $n+1$ $x = 0, 1, \cdots, n$ c) π d) $n\pi$

1 --- **6.** The proportion of times that a quantity x falls between the 5^{th} and the 90^{th} percentiles of x is
a) 85%, for any distribution   .05   .85   c) 85%, only if x is uniformly distributed.
b) 85%, only if x is normally distributed   d) 95%, for any distribution.
5th percentile   90th percentile

1 --- **7.** We observe the number of visits to a website per minute, x, on four different occasions and get 1, 3, 2, 2. Suppose x follows the Poisson distribution with parameter $\lambda = 3$. In the long run, what is proportion of minutes during which we expect to find the average number of visits per minute?
a) 1/4 $\left.\frac{e^{-\lambda}\lambda^x}{x!}\right|_{x=\lambda=3}$ b) 2/4 → $= e^{-3}3^3/3!$ c) 8/4   d) none of the above.

1 how 7, 7, **8.** The proportion of times when $X = \mu_x$ for the following distributions is (Circle "zero" or "nonzero")
ignore!
(*) a) Binomial(n, π)   zero/nonzero   c) Normal($\mu = 10, \sigma = 1$)   zero/nonzero
b) Poisson(λ)   zero/nonzero   d) Exponential(λ)   zero/nonzero

1 **9.** For the histogram shown here,
a) The mean (i.e., "location") and the std. dev. (i.e., "width") can be computed.
b) Only the mean (i.e., location) can be computed.
c) Only the standard deviation (i.e., width) can be computed.
d) Neither can be computed.

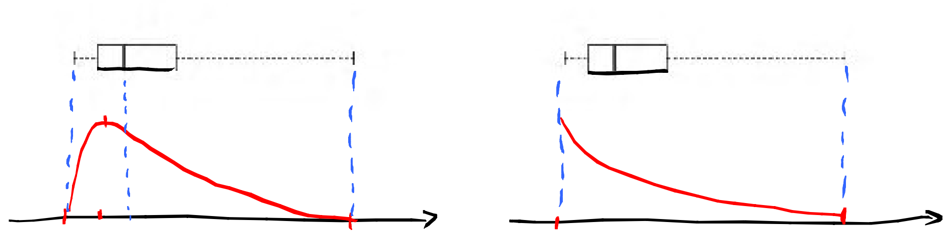


Just because a hist. has a weird "shape", it doesn't mean you cannot compute mean and std. dev.

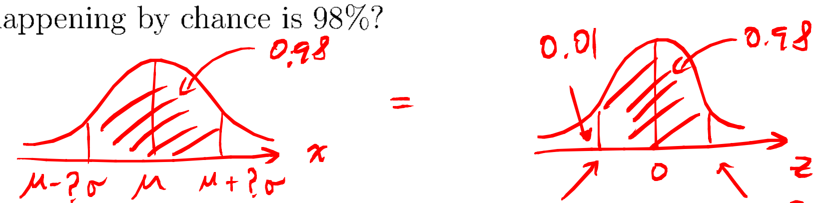
~~~2~~ 1.5 lect4,7 **10.** Treating the horizontal line under each boxplot as the x-axis representing the random values taken by a continuous random variable, draw the shape of two possible/likely distributions for the following boxplot. Make sure your drawings are as accurate as possible, taking into account everything you know.

(*) I had intended The answers to a) and b) to be nonzero, because prop(X=x) is non-zero for integer x. However, one student pointed out That μ_x may be real,
~~~

In which case prop(x = real) = 0 (zero). So, I have not graded parts a) and b).



~~2~~ 2.5 **11.** Physicists often talk about the 2-sigma rule, wherein any x within 2σ of μ of a normal distribution is considered to happen by chance. If they want an even more certain result, they follow the 3-sigma rule, etc. up to 5-sigma. What kind of a sigma rule is required so that the chance of x happening by chance is 98%?



$\frac{\mu-?\sigma-\mu}{\sigma}=-?$ $\quad \frac{\mu+?\sigma-\mu}{\sigma}=+?$

Table 1: $\boxed{?=2.33}$ $\Leftarrow$

$\boxed{2.33\ \sigma \text{ rule}}$

12. For simplicity, assume that there is a large number of cities on Earth (say 10^6) and that each city has a large number of buildings (say 10^3). Also, suppose that the proportion of defective buildings on Earth is very small, and that the average number of defective buildings per city is 10. Provide TWO ways of computing the proportion of cities that have 4 defective buildings per city. Show work, but the answer may be an expression with numbers; don't waste time on arithmetic.

~ 2 Way 1)

Since $n=10^3$ is "large", and prop. of defective π is very small

We can use Poisson with $\lambda = 10$ (avg # of defective building)

Then $P(x) = \frac{e^{-\lambda}\lambda^x}{x!} \Rightarrow \boxed{P(x=4) = \frac{e^{-10}10^4}{4!}}$

~ 2 Way 2)

In This case we are given n and $\lambda = n\pi$, so $\pi = \frac{\lambda}{n} = \frac{10}{10^3} = 0.01$

Then Binomial tells us

$p(x) = \binom{n}{x}\pi^x(1-\pi)^{n-x} = \boxed{\binom{10^3}{4}(.01)^4(.99)^{100-4}}$

~~2~~ 1.5 Q&A **13.** If we are interested in a Poisson variable with $\lambda = 4$ being within 4.8 of the mean, what values of the variable are we interested in?

For Poisson $\mu_x = \lambda = 4$

"within 4.8 of The mean" $= \mu_x - 4.8 < x < \mu_x + 4.8$

$-0.8 < x < 8.8$

For Poisson x = integer $\boxed{x = 0, 1, 2, \ldots, 8}$

Sample tests

~ 2 **14.** Find the n^{th} percentile of $f(x) = 2x, \quad 0 < x < 1$.

$$\int_{-\infty}^{n^{th}\text{ percentile}} f(x)\,dx = \frac{n}{100} \Rightarrow \int_0^{n^{th}\text{ percentile}} 2x\,dx = \frac{n}{100} \Rightarrow 2\frac{1}{2}x^2\Big|_0^{n^{th}\text{ perc.}} = \frac{n}{100}$$

$$\boxed{n^{th}\text{ percentile} = \frac{\sqrt{n}}{10}} \Leftarrow (n^{th}\text{ perc.})^2 = \frac{n}{100}$$

$\sim \cancel{2}$ 1.5 **15.** Find the mean of the Bernoulli distribution $p(x) = \pi^x(1-\pi)^{1-x}, \quad x = 0, 1$

$$\mu_x = \sum_x x\,p(x) = \sum_{x=0,1} x\,\pi^x(1-\pi)^{1-x} = 0\,\pi^0(1-\pi)^1 + 1\,\pi^1(1-\pi)^0$$

$$\boxed{\mu_x = \pi}$$

$\sim \cancel{2}$ 2.5 **16. Starting from the computational formula for the sample variance (and NOT rewriting it in terms of the defining formula)**, show how it changes if all data are shifted by a positive constant c.

$$s^2 = \frac{n}{n-1}\left(\overline{x^2} - \bar{x}^2\right) = \frac{n}{n-1}\left[\frac{1}{n}\sum_i x_i^2 - \left(\frac{1}{n}\sum_i x_i\right)^2\right]$$

$$x_i \to x_i + c$$

$$s^2 \to \frac{n}{n-1}\Big[\underbrace{\frac{1}{n}\sum_i (x_i+c)^2}_{\frac{1}{n}\sum x_i^2 + \frac{2c}{n}\sum x_i + \frac{c^2}{\cancel{n}}\cancel{\sum_i 1} \;=\; \overline{x^2} + 2c\bar{x} + c^2} - \Big(\underbrace{\frac{1}{n}\sum (x_i+c)}_{\frac{1}{n}\sum x_i + \frac{c}{\cancel{n}}\cancel{\sum 1} \;=\; \bar{x} + c}\Big)^2\Big]$$

$$s^2 \to \frac{n}{n-1}\left[\overline{x^2} + \cancel{2c\bar{x}} + \cancel{c^2} - \bar{x}^2 - \cancel{2c\bar{x}} - \cancel{c^2}\right]$$

$$\to \frac{n}{n-1}\left(\overline{x^2} - \bar{x}^2\right)$$

$s^2 \to s^2$ (s^2 does not change).