

QUIZ #1

Due Monday, March 17, 2003 by 4:00 pm

Name: SOLUTIONS

For each of the following questions, please be sure to (1) **show your work**, and (2) place a **box around your final answer** (except for entries in tables). If you need to make substantial assumptions, be sure to state them.

Question 1

(a) Consider a Poisson random variable y with parameter $\lambda = 3$, i.e., $y \sim P(\lambda = 3)$. If we take the sum of 5 random samples (independently and identically distributed) of y , how would the sum $x = y_1 + y_2 + \dots + y_5$ be distributed? (Note that when we ask how a random variable is distributed, we need both the form and the parameters of that distribution.)

$$E(x) = E(y_1) + E(y_2) + \dots + E(y_5)$$

$$= 3 + 3 + \dots + 3 = 15$$

5 pts

$x \sim P(15)$
 - Poisson with
 $\lambda = 15$

$$\text{Var}(x) = \text{Var}(y_1) + \text{Var}(y_2) + \dots + \text{Var}(y_5) = 15$$

A reasonable approximation (by CLT) is $x \sim N(15, 15)$

(b) Consider the same random variable $y \sim P(\lambda = 3)$. If we calculate \bar{x} as the average of y_1, y_2, \dots, y_5 , how is \bar{x} distributed?

$$E(\bar{x}) = \frac{1}{5} \{ E(y_1) + \dots + E(y_5) \} = 3$$

$$\text{Var}(\bar{x}) = \text{Var}\left(\frac{y_1}{5}\right) + \dots + \text{Var}\left(\frac{y_5}{5}\right)$$

$$= \frac{1}{25} \{ \text{Var}(y_1) + \dots + \text{Var}(y_5) \} = \frac{15}{25} = \frac{3}{5}$$

5 pts

Now the average (as opposed to the sum) of Poisson random variables is approximately Normal: $\bar{x} \sim N(3, 3/5)$

(c) Now suppose that $y \sim N(0, 1)$ (zero mean in this case). If we calculate x as the sum of the squares of the samples, how is $x = y_1^2 + y_2^2 + \dots + y_5^2$ distributed? What is the expected value of x , $E(x)$?

$x \sim \chi_5^2$
 $E(x) = 5$

by definition of chi-square

5 pts

Question 2

We are interested in a control chart to monitor particle defects at an important point in the process. We will sample one wafer from each lot processed on the tool, and count the number of defects observed on that wafer. Preliminary data is shown in Table 1.

Table 1: Defects observed on sample wafer, preliminary runs

Lot #	1	2	3	4	5	6	7	8	9	10
# Defects	16	14	15	12	15	19	15	16	15	13

(a) Assuming the number of defects x on a wafer is Poisson distributed, estimate μ_x and σ_x for the mean and standard deviation, respectively, of the number of defects on a wafer.

$$\bar{c} = \hat{\mu}_x = \frac{\sum d_i}{n} = \frac{150}{10} = 15$$

$$\hat{\sigma}_x = \sqrt{\bar{c}} = \sqrt{15} = 3.87$$

Since Poisson, we know

$$\lambda = \mu_x = \sigma_x^2$$

5 pts

(b) We are very concerned about getting this process under control, and are willing to investigate possible significant events at a fairly high rate (or short average run length for false alarms). We thus set up a conventional control chart, but with $2\text{-}\sigma$ control limits, i.e. $UCL = \mu_x + 2\sigma_x$, and $LCL = \mu_x - 2\sigma_x$. Given that the process is Poisson distributed and actually is under statistical control, what is the probability of an observation lying above the upper control limit? What is the probability of an observation lying below the lower control limit? What is the average run length for "false alarms" using this control chart?

$$UCL = 15 + 2(3.87) = 22.75$$

$$LCL = 15 - 2(3.87) = 7.26$$

5 pts

Above; UCL: $Pr(x \geq 23) = 1 - Pr(x \leq 22) = 0.033$ from stat charts for $\lambda = 15$

Below; LCL: $Pr(x \leq 7) = 0.018$

$$ARL = \frac{1}{\alpha} = \frac{1}{(0.033 + 0.018)} = 19.6$$

(c) For comparison, what would the false alarm or α -risk and average run length be for a $\pm 2\sigma$ control chart such as that used above, if the sample values were actually normally distributed?

$$\alpha = 2(1 - \Phi(2)) \quad \text{using cumulative normal stats chart}$$

$$= 2(1 - 0.97725)$$

$$\alpha = 0.0455$$

5 pts

$$ARL_0 = \frac{1}{\alpha} = 21.98$$

Question 3

The process variance of the thickness of a CVD tantalum nitride barrier layer has been a problem. Historical data (based on hundreds of runs) has indicated that the variance has been 10 \AA^2 . Recently, a new wafer chuck has been provided by the tool vendor, with a claim that it will reduce variance by a factor of 2.5 (i.e. to a variance of 4 \AA^2). You want to run an experiment to see if there is any evidence that the new chuck does not in fact reduce the variance by this amount. You will conduct n new runs on the tool using the new chuck, and decide that if you observe a new variance of 8 \AA^2 or higher, you will declare that the variance has not in fact been reduced, to a confidence of 95%. How many runs (what value of n) will you need in order to make this statement?

$$\sigma_{\text{old}}^2 = 10 \quad \text{historical variance}$$

15 pts

$$\sigma_{\text{new}}^2 = 4 \quad \text{claimed (hypothesized) new variance.}$$

If the true new variance is 4, then by sampling alone we would expect S_{new}^2 to vary somewhat. So the idea is to use the F distribution to help us understand the likelihood (and control that likelihood to 0.05) of observing $S_{\text{new}}^2 = 8$ if the new variance really is 4.

$$\frac{\sigma_{\text{OLD}}^2}{\sigma_{\text{NEW}}^2} \leq F_{\alpha, \nu_{\text{NEW}}, \nu_{\text{OLD}}} \frac{S_{\text{OLD}}^2}{S_{\text{NEW}}^2}$$

But $S_{\text{OLD}}^2 = \sigma_{\text{OLD}}^2$ based on "infinite" existing samples. So

$$\frac{\cancel{\sigma_{\text{OLD}}^2}}{\sigma_{\text{NEW}}^2} \leq F_{0.05, \nu_{\text{NEW}}, \infty} \frac{\cancel{\sigma_{\text{OLD}}^2}}{S_{\text{NEW}}^2}$$

$$2 = \frac{8}{4} = \frac{S_{\text{new}}^2}{\sigma_{\text{new}}^2} \leq F_{0.05, \nu_{\text{NEW}}, \infty}$$

So use $F_{0.05}$ charts with $\nu_1, \nu_2 = \infty$ to find a value ≥ 2 .
We find $F_{0.05, 7, \infty} = 2.01$

$\therefore n = 8$ samples needed to reject vendor's claim with 95% confidence if we observe $S_{\text{new}}^2 \geq 8$.

Question 3, continued

- An alternative approach is perhaps a little clearer, and makes clear that this is purely a hypothesis test on the new distribution.

Here we are assuming the vendor's new process is working ($\sigma_{\text{NEW}}^2 = 4A^2$), and are looking for evidence this is not true... i.e. that variance is bigger than this.

So, we have a one-sided confidence interval

$$\sigma_{\text{new}}^2 \geq \frac{(n-1) S_{\text{new}}^2}{\chi^2_{\alpha, n-1}} \quad \text{when } \nu = n-1$$

or

$$\chi^2_{0.05, \nu} \geq \nu \frac{S_{\text{new}}^2}{\sigma_{\text{new}}^2} = \nu \cdot \frac{8A^2}{4A^2} = 2\nu$$

\swarrow decision point
 \nwarrow assumed variance

So now we can use χ^2 chart under the $\alpha = 0.05$ column. Unfortunately we have ν both in $\chi^2_{0.05, \nu}$ and in 2ν ... so try iteratively

ν	$\chi^2_{0.05, \nu}$	\geq	2ν	
15	25	\nless	30	NO
10	18.31	\nless	20	NO
5	11.67	\nless	10	YES
6	12.59	\nless	12	YES
7	14.67	\nless	14	YES
8	15.51	\nless	16	NO

... look for largest ν for which this is true

$\Rightarrow \nu = 7$ or $n-1 = 7 \Rightarrow n = 8$

Same as with F distribution.

Question 4

We wish to design an \bar{x} control chart to monitor the capacitance of a large patterned metal-insulator-metal (MIM) interconnect capacitor. We are fortunate because the process is highly capable. To be specific, the process is normally distributed, the mean is 100, the process standard deviation is 10, and the upper specification limit is 170.

(a) The process mean is known to wander somewhat, but is generally between 100 and 110. It never wanders or shifts below 100. What is the best case process capability? What is the worst case process capability?

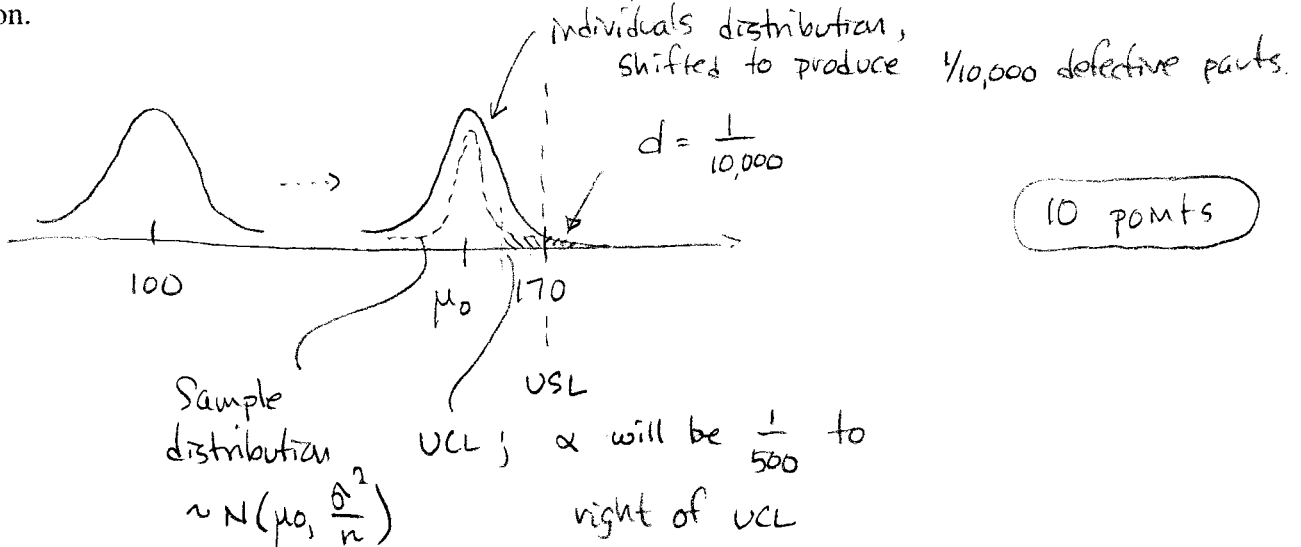
$$C_{pk} = \min \left\{ \frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma} \right\}$$

5 points

Best case: $\mu = 100$: $C_{pk} = \frac{170 - 100}{3 \cdot 10} = \frac{7}{3} = 2.33$

Worst case: $\mu = 110$: $C_{pk} = \frac{170 - 110}{3 \cdot 10} = \frac{6}{3} = 2$

(b) We will only be concerned with the process if the mean drifts so much toward the upper specification limit that it starts to produce defective parts at a rate greater than one per every 10,000 parts. We will be using an \bar{x} chart with sample size $n = 20$. What should our upper control limit (UCL) be if we are willing to tolerate false alarms at a rate of 1 per 500 runs? Carefully explain your reasoning in all steps of your solution.



- First, the defective rate of $d = 0.0001$ sets where μ_0 needs to be for us to worry:

$$\Phi\left(\frac{170 - \mu_0}{10}\right) = 0.9999 \quad \text{gives } z = 3.71 \text{ from cumulat. std. normal chart}$$

- False alarm rate of $\alpha = 1/500 = 0.002$ occurs for $1 - \Phi(z) = 0.002$

- Since we're sampling w/ $n = 20$, the UCL is $\Phi(z) = 0.998 \Rightarrow z = 2.88$ by chart
- $$UCL = \mu_0 + z_\alpha \cdot \frac{\sigma}{\sqrt{n}} = 132.9 + 2.88 \cdot \frac{10}{\sqrt{20}} = \boxed{139.3 = UCL}$$