STAT 113 – Midterm 2 SOLUTION

1a. [1pt] Suppose y is a normally distributed random variable with mean 0 and variance 1.0, i.e. y is standard normal. Find $P(-1.0 < y < 0.5)$.

$P(-1.0 < y < 0.5) = P(0 < z < 1) + P(0 < z < 0.5) = 0.34 + 0.19 = 0.53.$

1b. [2pt] Suppose y is normally distributed random variable with mean 10 and variance 2^2, i.e., $y \sim N(0, 2^2)$. Find the value y_0, such that $P(y > y_0) = 0.25$.

Let z denote the standard normal random variable $z = \frac{y-\mu}{\sigma}$.
Then $P(z > z_0) = 0.1$ for $z_0 = 1.28$, and thus $P(y > y_0) = 0.1$ for $y_0 = \mu + \sigma z_0 = 12.56$.

1c. [2pt] Assume x and y are independent random variables with $E(x) = 1.0$, $E(y) = 10.0$ and $Var(x) = 1.0$, $Var(y) = 4.0$. Let $z = 2x - y$. Find $E(z)$ and $Var(z)$.

$E(z) = 2E(x) - E(y) = 2 - 10 = -8$ and
$Var(z) = 4 \cdot Var(x) + 1 \cdot Var(y) = 4 + 4 = 8.$

1d. [1pt] Suppose x is a χ^2 distributed random variable with $\nu = 10$ degress of freedom. Find the value x_0, such that $P(x > x_0) = 0.100$.

Using the table on page 1100, line 10, find: $x_0 = 15.9871$.

1e. [4pt] A chromosome mutation believed to be linked with color blindness is knwon to occur, on the average, once in every 10,000 births. If 20,000 babies are born this year in a certain city, using the Poisson approximation to the Binomial distribution find the probabilty that exactly three of the children have the mutation.

With $\lambda = 2$, we find $p(3) = \frac{\lambda^3 e^{-\lambda}}{3!} = 0.18$.

2. Let c be a constant and consider the density function

$$f(y) = \begin{cases} (1/c)e^{-y/2} & \text{if } y \geq 0 \\ (1/c)e^{y/2} & \text{if } y < 0 \end{cases}$$

2a. [4pt] Find the value of c.

Note the symmetry around the origin $y = 0$.

$$\frac{1}{c}\int_0^\infty e^{-y/2}dy = \frac{1}{2}$$

$\Rightarrow c = 4.$

2b. [3pt] Find the cumulative distribution function $F(y)$.
Hint: Consider two cases: $y < 0$ and $y \geq 0$.

For $y \leq 0$: $F(y) = \frac{1}{4}\int_{-\infty}^{y} e^{y/2}dy = \frac{1}{2}e^{y/2}$
For $y > 0$: $F(y) = \frac{1}{2} + \frac{1}{4}\int_0^y e^{-y/2}dy = 1 - \frac{1}{2}e^{-y/2}$.

2c. [1pt] Compute $F(1)$.

$F(1) = 1 - \frac{1}{2}e^{-1/2} = 0.70$

2d. [1pt] Find $P(y > 0.5)$.

$P(y > 0.5) = 1 - F(0.5) = \frac{1}{2}e^{-1/4} = 0.39.$

2e. [1pt] Find $P(y > 0.5 | y > 0)$.

$P(y > 0.5 | y > 0) = \frac{P(y>0.5)}{P(y>0)} = 2 \cdot 0.39 = 0.78.$

3. y_i, $i = 1, \ldots, n$ with $\mu = 9.2$, $\sigma^2 = 4.0$, $n = 100$.

3a. [2pt] Denote with $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ the sample mean. Assuming the compensation program was *not* effective in reducing the average number of sick days used, find $E(\bar{y})$ and $Var(\bar{y})$.

$E(\bar{y}) = \mu = 9.2$ and
$Var(\bar{y}) = \sigma^2/n = 4.0/100 = 0.04$.

3b. [4pt] Find a values c, such that $P(\bar{y} - c < \mu < \bar{y} + c) = 0.95$.

By the central limit theorem, $\bar{y} \sim N(\mu, \sigma^2/n)$ and thus $z = \frac{\bar{y}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.
Therefore

$$P(-1.96 < \underbrace{\frac{\bar{y}-\mu}{\sigma/\sqrt{n}}}_{z} < 1.96) = 0.95$$

$$P(\bar{y} - 1.96 \cdot \sigma/\sqrt{n} < \mu < \bar{y} + 1.96 \cdot \sigma/\sqrt{n}) = 0.95$$

i.e., $c = 1.96 \cdot \sigma/\sqrt{n} = 1.962.0/10.0 = 0.392$.

3c. [4pt] Assuming the compensation program was *not* effective in reducing the average number of sick days used, find the probability that $\bar{y}$, the mean number of sick days used by the sample of 100 managers, is leass than 8.80 days, i.e., find $P(\bar{y} < 8.80)$.

$$P(\bar{y} < 8.8) = P(\underbrace{\frac{\bar{y}-\mu}{\sigma/\sqrt{n}}}_{z} < \underbrace{\frac{8.8-\mu}{\sigma/\sqrt{n}}}_{-1.0}) = \frac{1}{2} - P(0 < z < 1.0) = \frac{1}{2} - 0.3412 = 0.16.$$

4. Sixty six bulk specimens of Chilean lumpy iron ore were randomly sampled from a shipload of ore, and the percentage of iron in each ore specimen was determined. The data y_i, $i = 1, \ldots, 66$, are shown in the following table

4a. [2pt] Using MINITAB (or any other statistics program of your choice) construct a relative frequency histogram for the data (*cut and paste the printout here*).

```
61.6     1  *
61.8     3  ***
62.0     6  ******
62.2     1  *
62.4     4  ****
62.6     3  ***
62.8    10  **********
63.0    12  ************
63.2     9  *********
63.4     6  ******
63.6     4  ****
63.8     2  **
64.0     4  ****
64.2     0
64.4     1  *
```

4b. [2pt] Find sample mean $\bar{y}$ and sample variance s^2.

$\bar{x} = 62.963$, $s^2 = 0.609^2$.

4c. [3pt] Let $\mu = E(y_i)$ and $\sigma^2 = Var(y_i)$ denote expected value and variance of the measurements. What is the sampling distribution of the statistic $\frac{\bar{y}-\mu}{s/\sqrt{n}}$?

$t = \frac{\bar{y}-\mu}{s/\sqrt{n}} \sim t$ with $\nu = n - 1$ d.f.

4d. [3pt] Find a value c, such that $P(\bar{y} - c < \mu < \bar{y} + c) = 0.95$.

Find from Table, on page 1099:

$$P(-2.0 < \underbrace{\frac{\bar{y}-\mu}{s/\sqrt{n}}}_{t} < 2.0) = 0.95$$

and therefore

$$P(\underbrace{\bar{y} - 2.0s/sqrtn}_{=63.0+2\cdot0.61/\sqrt{66}} < \mu < \bar{y} + 2.0s/sqrtn) = 0.95$$

i.e., $P(62.85 < \mu < 63.15) = 0.95$.