Answer 7.26

The hypothesis is given by: $H_0: p = p_0 = 0.04$, $H_1: p > 0.04$.

Since $np_0q_0 = 50 \times 0.04 \times 0.96 = 1.92 < 5$, we will use the exact test.

$$p = \sum_{k=5}^{50} {}_{50}C_k\,(0.04)^k(0.96)^{50-k} = 1 - \sum_{k=0}^{4} C_k(0.04)^k 0.96^{50-k}$$

$$\Pr(X \leq 4|n = 50, p = 0.04) = 0.9510$$

Answer 7.30

Let p= true proportion of death lung cancer for workers in the plant. The hypothesis is:

$H_0: p = 0.12$, $H_1: p \neq 0.12$.

Answer 7.31

A two-sided test is appropriate as there is no prior information to determine whether the proportion of lung cancer deaths is higher or lower than in the general population.

Answer 7.32

As $np_0q_0 = 20 \times 0.12 \times 0.88 = 2.1 < 5$, we will use the exact test.

Since $\hat{p} = 0.25 > 0.12$

$$p = 2 \times \sum_{k=5}^{20} C_k(0.12)^k 0.88^{20-k}$$

$$\Pr(X \leq 4|n = 20, p = 0.12) = 0.917$$

$$p = 2 \times (1 - 0.917) = 0.166$$

Answer 7.48

$$SMR = \frac{17}{6.3} = 270\%$$

Answer 7.49

Since the expectation of number of events is less than 10, we will use the exact test.

When $x = 17$, the 95% confidence level for mean is (9.90, 27.22). Therefore, we will reject H_0.

$$p = 2 \times \sum_{k=17}^{\infty} \frac{e^{-6.3}6.3^k}{k!} = 2 \times (1 - \sum_{k=0}^{16} \frac{e^{-6.3}6.3^k}{k!})$$

$$\Pr(X \leq 16|\mu = 6.3) = 0.9997$$

$$p = 2 \times (1 - 0.9997) < 0.001$$

Answer 7.109

No. A two-sided p-value of 0.03 means that we will reject the null hypothesis at 0.05 level. However, if the 95% CI contains the null value 2, we should not reject the null hypothesis at the 0.05 level.

Answer A1

The hypothesis is given by: $H_0: \mu = 400$, $H_1: \mu \neq 400$.

$$t = \frac{388 - 400}{\sqrt{\frac{37}{35}}} = -11.671$$

The t_{34} critical value at 0.05 significant level (two tail) is 2.032

Since 11.671>2.032, we will reject the null hypothesis and conclude that the intake of women whose income is below the poverty level differs from the recommended amount.

Answer A2

The sample size estimation is performed in STATA:

```
. power onemean 400 388, power(0.9) sd(37)

Performing iteration ...

Estimated sample size for a one-sample mean test
t test
H0: m = m0  versus  Ha: m != m0

Study parameters:

        alpha =    0.0500
        power =    0.9000
        delta =   -0.3243
           m0 =  400.0000
           ma =  388.0000
           sd =   37.0000

Estimated sample size:

            N =       102
```

Answer B1

The lighter smoking twin is labelled as twin 1. We are only interested in one variable, which is the density of the femoral neck bone fn1. A one-sample t test will be appropriate.

Answer B2

The hypothesis is given by $H_0: \mu = 0.77$, $H_1: \mu \neq 0.77$.

Answer B3

```
. ttest fn1 == 0.77

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
     fn1 |      41     .664878    .0162943    .1043341    .6319461      .69781
------------------------------------------------------------------------------
    mean = mean(fn1)                                              t =  -6.4515
H0: mean = 0.77                                  Degrees of freedom =       40

   Ha: mean < 0.77             Ha: mean != 0.77               Ha: mean > 0.77
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000

. display invttail(40, 0.025)
2.0210754
```

Since $t_{40} = 2.02 < 6.45$, we will reject the null hypothesis and conclude that the lighter smoking twins do not have a healthy bone density.

Answer B4

There is only one variable we interested in, which is the density of the femoral neck bone fn1. We will choose the one-sample test and test the normality of this variable.

```
. swilk fn1

                   Shapiro–Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+------------------------------------------------------
         fn1 |         41    0.98740      0.508    -1.428    0.92340
```

The result shows that this variable is normally distributed. The hypothesis is given by: $H_o: \mu = 0.77$, $H_1: \mu < 0.77$.

```
. ci means fn1, level(99)

    Variable |        Obs        Mean    Std. err.       [99% conf. interval]
-------------+---------------------------------------------------------------
         fn1 |         41     .664878    .0162943        .6208109    .7089452
```

Since the upper bound of confidence interval is 0.709 <0.77, we can reject the null hypothesis and conclude that we are 99% confident that the bone density of the lighter smoking twin lower than heathy levels.