**EXAMINATION:** Tentamensskrivning i Matematisk Statistik (TMS061)

*Time:* Onsdag 27 Maj 2009

*Aid:* You are allowed to use a scientific calculator and a one page (both sides) of notes. In the notes you are allowed to have only formulas and their names. If there are any definitions or theorems or more explanations included the notes will be taken away.

**Grade:** You need 34 out of (30+10) points for 5, 24 points for 4 and 17 points for 3.
You still fail the course if you get less than 12 points in the written exam!
**Motivate all your answers. Good Luck!**

**1**) A project manager has 10 chemical engineers on her staff. Four are women and six are men. These engineers are equally qualified. In a random selection of three workers, what is the probability that no women are selected? Would you consider it unusual for no women to be selected under these circumstances? Explain! (p)

**2**) Let $A$ and $B$ be two events.

a) Assume that $A$ and $B$ are mutually exclusive such that: $P(A)P(B) > 0$. Show that these events are not independent.

b) Let $P(A) = 0.5$ and $P(B) = 0.7$. What must be $P(A \cup B)$ equal to for $A$ and $B$ to be independent?

**3**) In studying the causes of power failures, these data have been gathered.
5% are due to transformer damage
80% are due to line damage
1% involve both problems
Based on these percentages, approximate the probability that a given power failure involves

a) line damage given that there is transformer damage

b) transformer damage but not line damage?

**4)** Let $X$ denote the time required to upgrade a computer system in hours. Assume that the density for $X$ is given by:

$$f(x) = ke^{-2x}, \quad 0 < x < \infty$$

a) Find the numerical value of $k$ that makes this a valid density.

b) Find the probability that it will take at most 1 hour to upgrade a given system.

**5)** Let $X_1, X_2, X_3, X_4, X_5$ be a random sample from a binomial distribution with $n = 10$ and $p$ unknown.

a) Verify that $\frac{\bar{X}}{10}$ is an unbiased estimator for $p$.

b) Estimate $p$ based on these data: $3, 4, 4, 5, 6$

c) Obtain the maximum likelihood estimator of $p$. Does this contradict your answer in (a)?

**6)** Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with mean $\mu$ and variance $\sigma^2$.
a) Explain why $\bar{X}$ is distributed as normal random variable. Compute $E(\bar{X})$ and $Var(\bar{X})$.

b) Let $Y = 2X_1 - 3X_2 - 5$ where $X_1, X_2$ are from the above random sample. Compute $E(Y)$ and $Var(Y)$.

**7)** Researchers are experimenting with a new compound used to bond Teflon to steel. The compounds currently in use require an average drying time of 3 minutes. It is thought that the new compound dries in a shorter length of time.
a) Set up the null and alternative hypotheses needed to support the claim that the new compound dries faster that the one currently in use.

b) Discuss the practical consequences of making a Type I error.

c) When the experiment is conducted, these data are obtained:

| | | | |
|---|---|---|---|
| 1.4 | 2.1 | 2.8 | 0.9 |
| 2.4 | 1.7 | 3.7 | 2.7 |
| 2.6 | 1.9 | 2.8 | 2.8 |
| 2.2 | 2.2 | 3.4 | 1.9 |

with $\bar{x} = 2.3438, s^2 = 0.5106$. Test the null hypothesis of part (a) at $\alpha = 0.05$ level. Would you suggest of marketing this new product?

d) Repeat part (c) only now assuming that from a pilot study you know that $\sigma^2 = 0.5106$.

**8** In a study of the association between color and the effectiveness of a graphical display 100 graphs are randomly selected from among current scientific journals. Each is classified as to whether or not color is used. Each is also rated as to its effectiveness in making its point. Resulting data are given in the next table:

Color present

| Effective | Yes | No |
|---|---|---|
| Excellent | 7 | 4 |
| Good | 10 | 19 |
| Fair | 9 | 26 |
| Poor | 4 | 21 |

Is there evidence that the effectiveness of a graphical display is not independent of color? Explain. Use $\alpha = 0.05$.

**9** a) The estimator of the true mean of the dependent variable $Y$ was given by $\hat{\mu}_{Y|x} = \beta_0 + \beta_1 x$. Show that $E(\hat{\mu}_{Y|x}) = \mu_{Y|x}$, and hence $\hat{\mu}_{Y|x}$ is an unbiased estimator for $\mu_{Y|x}$.

b) Show that for a simple linear regression model $\sum_{i=1}^{n}(y_i - \hat{y}_i) = 0$