## Key to Fall 2004 Final Exam

1. (a) $E(Z) = E(3X - 2Y) = E(3X) - E(2Y) = 3E(X) - 2E(Y) = 3(2) - 2(2) = 2$

   (b) $\text{Var}(W) = \text{Var}(3X - 2Y + 5) = \text{Var}(3X) + \text{Var}(2Y) = 3^2\text{Var}(X) + 2^2\text{Var}(Y)$
   $= 9(4) + 4(9) = 72$

   (c) $\text{Var}(Z) = \text{Var}(3X - 2Y) = 3^2\text{Var}(X) + (-2)^2\text{Var}(Y) = 9(4) + 4(9) = 72$. Linear combinations of normal random variables are normally distributed, so
   $Z \sim N(\mu = 2, \sigma^2 = 72)$

2. (a) $f_X(x) = \begin{cases} 0.4 & x = 1 \\ 0.6 & x = 2 \\ 0 & \text{o.w.} \end{cases}$

   (b) $E(X) = 1(.4) + 2(.6) = 1.6$

   (c) $E(X^2) = 1(.4) + 4(.6) = 2.8$, $\text{Var}(X) = 2.8 - 1.6^2 = 2.8 - 2.56 = 0.24$

   (d) $P(Y = 1|X = 2) = \frac{f_{XY}(2,1)}{f_X(2)} = \frac{0.45}{0.6} = 0.75$

   (e) $E(XY) = (1)(1)(0.3)+(1)(2)(0.1)+(2)(1)(0.45)+(2)(2)(0.15) = 0.3+0.2+0.9+0.6 = 2.0$

   (f) $E(Y) = 1(0.75) + 2(0.25) = 1.25$, so $\text{Cov}(X, Y) = E(XY) - (E(X))(E(Y)) = 2 - (1.6)(1.25) = 2 - 2 = 0$

   (g) $\text{Corr}(X, Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(y)}} = \frac{0}{\sqrt{\text{Var}(X)\text{Var}(y)}} = 0$

   (h) Since the correlation is 0 the best linear relationship is a horizontal line, i.e. there is neither a positive nor a negative linear relationship between $X$ and $Y$. This is the weakest possible linear relationship.

3. (a) $P(X < 2.3) = P(\frac{X-3}{0.5} < \frac{2.3-3}{0.5}) = P(Z < -1.40) = 0.0808$

   (b) $P(2.25 < X < 3.75) = P(\frac{2.25-3}{0.5} < \frac{X-3}{0.5} < \frac{3.75-3}{0.5}) = P(-1.50 < Z < 1.5) = 0.9332 - 0.0668 = 0.8664$

   (c) $\frac{x-3}{0.5} = -1.04$, so $x = 3 - 0.52 = 2.48$

4. (a) Fixed number of Bernoulli trials, independent with $p = 0.8$ so
   $X \sim \text{Binomial}(n = 100, p = 0.8)$.

   (b) $E(X) = 100(.8) = 80$, $\text{Var}(X) = 100(.8)(.2) = 16$

   (c) $P(X \geq 75) = P(\frac{X-80}{\sqrt{16}} \geq \frac{75-80}{\sqrt{16}}) \doteq P(Z \geq -1.25) = .8944$

5. (a) $E(X) = \mu = 4(.2) + 5(.4) + 6(.3) + 7(.1) = 5.3$
   $E(X^2) = 16(.2) + 25(.4) + 36(.3) + 49(.1) = 28.9$
   $\text{Var}(X) = \sigma^2 = 28.9 - 5.3^2 = 28.90 - 28.09 = 0.81$

   (b) $\mu_{\bar{X}} = \mu = 5.3$
   $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} = \frac{0.81}{36} = 0.0225$

   (c) Since $n$ is large (over 30 – note that this works here but not necessarily for the binomial) then $\bar{X}$ is approximately normally distributed with $\mu_{\bar{X}} = 5.3$ and $\sigma^2_{\bar{X}} = 0.0225$

(d) $P(\bar{X} < 5.5) = P(\frac{\bar{X}-5.3}{\sqrt{0.0225}} < \frac{5.5-5.3}{\sqrt{0.0225}}) \doteq P(Z < 1.33) = 0.9082$

6. (a) $\bar{x} \pm t_{.025} s/\sqrt{n}$ is $325.05 \pm 2.064(0.5)/\sqrt{25}$ or $325.05 \pm 0.206$, i.e. $324.844 \leq \mu \leq 325.256$

   (b) We do not know what $\mu$ is for the population of contents in mg for bottles of buffered aspirin. In order to have obtained the data we did, however, $\mu$ must be somewhere between 324.844 mg and 325.256 mg. We are correct 95% of the time when we make claims like this, i.e. only 5% of the time intervals we construct in this way do not actually contain $\mu$.

7. (a) $1 = \int_0^2 cx(1+x)dx = c(\frac{x^2}{2} + \frac{x^3}{3})|_0^2 = c(2 + \frac{8}{3}) = c\frac{14}{3} \Rightarrow c = \frac{3}{14}$

   (b) $P(X > 1) = \int_1^2 \frac{3}{14}x(1-x)dx = \frac{3}{14}(\frac{x^2}{2} + \frac{x^3}{3})|_1^2 = \frac{3}{14}(\frac{14}{3} - \frac{5}{6}) = 1 - \frac{5}{28} = \frac{23}{28} = 0.8214$

   (c) $E(X^2+2) = E(X^2) + 2 = \int_0^2 \frac{3}{14}x^2 x(1+x)dx + 2 = \int_0^2 \frac{3}{14}x^3(1+x)dx + 2 = \frac{3}{14}(\frac{x^4}{4} + \frac{x^5}{5})|_0^2 + 2 = \frac{3}{14}(4 + \frac{32}{5}) + 2 = \frac{3}{14}(\frac{52}{5}) + 2 = \frac{78}{35} + 2 = \frac{148}{35} \doteq 4.23$

   (d) $\int_0^x \frac{3}{14}w(1+w)dw = \frac{3}{14}(\frac{w^2}{2} + \frac{w^3}{3})|_0^x = \frac{3}{14}(\frac{x^2}{2} + \frac{x^3}{3}) = \frac{3x^2+2x^3}{28}$ so

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{3x^2+2x^3}{28} & 0 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

8. (a) $\hat{p} = \frac{1066}{1600} = 0.66625$
   $\hat{p} \pm z_{.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.66625 \pm 1.96\sqrt{\frac{0.66625(0.33375)}{1600}} \doteq 0.666 \pm 0.023$, or $0.643 \leq p \leq 0.689$

   (b) The interval does not include 0.75, so it is not plausible that we could have sampled from a population with $p = 0.75$ and obtained the data we got. The hypothesis that 75% of adults agree with the issue is not supported by the interval.