Answers of Master's Exam, Spring, 2009

1) a) $\dfrac{13!13!2^{13}}{26!}$ b) $\dfrac{13!2^{13}}{26!}$

2 a) 3/40 b) 2/3 c) 0.4332 d) 0.629 (normal approximation with ½ correction)

3) a) $\lim_{n\to\infty} F_n(x)$ = F(x) for every x on the real line at which F is continuous, where F is the c.d.f. of X.

b) For every $\varepsilon > 0$, $\lim_{n\to\infty} P(|Y_n - 5| > \varepsilon) = 0$.

c) $F(x) = e^x$ for $-\infty < x \le 0$, 1 for $x > 0$. F is the cdf of X.

4) a) $f_X(x) = 3x^2$ for $0 \le x \le 1$, 0 otherwise

b) $f_Y(y) = y^2$ for $0 \le y \le 1$,
$= y(2 - y)$ for $1 < y \le 2$
$= 0$ otherwise

c) $f_{Y|X}(y \mid x) = 2y/(3x^2)$ for $x \le y \le 2x$, $0 < x \le 1$

d) $E(Y \mid X = x) = (7/6)x$, for $0 < x \le 1$, so $Z = (7/6)X$

e) $E(Y|X = x) = E(Y|X=x)/(1+x) + x = (7/6)x/(1+x) + x$ for $0 < x \le 1$, so
$V = (7/6)X/(1+X) + X$

5) a) $U^{1/3}$ b) $f_M(m) = m^6$ for $0 \le m \le 1$, 0 for $m < 0$, 1 for $m > 1$.

c) $F_S(s) = (1/3)(1 + s)^3$ for $-1 \le s \le 0$, $s + (1 - s)^3/3$ for $0 < s \le 1$, 0 for $s < -1$,
1 for $s > 1$
$f_S(s) = (1 + s)^2$ for $-1 \le s \le 0$, $1 - (1 - s)^2$ for $0 < s \le 1$, 0 otherwise.

d) Let $Y = (X_3 - 1/3)^2$. Then
$F_Y(y) = (y^{1/2} + 1/3)^3 - (1/3 - y^{1/2})^3$ for $0 \le y \le 1/9$, $(y^{1/2} + 1/3)^3$ for $1/9 < y \le 4/9$,
0 for $y < 0$, 1 for $y > 4/9$.
$f_Y(y) = (3/2)(y^{1/2} + 1/3)^2 y^{-1/2} + (1/3 - y^{1/2})^2 y^{-1/2}$ for $0 \le y \le 1/9$
$= (3/2)(y^{1/2} + 1/3)^2 y^{-1/2}$ for $1/9 < y \le 4/9$, 0 otherwise.

e) $Var(X_1) = \sigma_1^2 = 1/12$, $Var(X_2) = \sigma_2^2 = 1/18$, $Var(X_3) = \sigma_3^2 = 3/80$
Let $Y_1 = 3X_1 - 7X_2 + 2X_3$, $Y_2 = -4X_1 + 5X_2 + 9X_3$.
$Cov(Y_1, Y_2) = -1.9319$, $Var(Y_1) = 3.8097$, $Var(Y_2) = 5.7597$,
$\rho(Y_1, Y_2) = -0.4124$

f) $E(X_3^2) = 3/5$, $E(X_1^{1/2}) = 2/3$, $E(1/(2 + X_2)) = 2(1- 2 \log(3) + 2 \log(2)) = 0.37814$, so E(W) = (product of these) = 0.15126.

Statistics

6) a) $\bar{X}/2$ b) $L(\sigma) = (1/\sigma)^n \exp(n – S/\sigma)$, where $S = \Sigma x_i$.. The plot has a concave shape with maximum at $\sigma = (1/n) \Sigma x_i$.   c) $\bar{X}$ d) Bias = σ.

7)   a) Reject for $X \leq \alpha^{1/4}$,   b) Neyman-Pearson,   c) $\alpha^{1/2}$

8) Let $X_1, \dots , X_7$ be the weights gains for the 7 cages with the strong magnetic field. Let $Y_1, \dots , Y_7$ be the weight gains for the 7 cages for the control cages. Suppose that $Y_1, \dots , Y_7$ is a random sample from a distribution with cdf F_X and $X_1, \dots , X_7$ is a random sample from a distribution with cdf F_Y, and these two sample are independent. The ranks to the X-sample are 7, 1, 6, 13, 4, 2, 3.
We wish to test H_0: $F_X = F_Y$ vs H_a: H_0 not true.
The Wilcoxon statistic is 7 + 1+ 6 + 13 + 4 + 2 + 3 = 36. The observed p-value is $2 P_0(W \leq 36) \doteq 2 \Phi((36.5-52.5)/ (49(15)/12)^{1/2}))$
$= 2 \Phi(-2.044) = 0.0409$. Reject H_0 are the $\alpha = 0.05$ level.

9) a) Let D_i = (Older Twin IQ) – (Younger Twin IQ) for pair i, i = 1, 2, … , 7. Suppose that the D_i are independent, each with a $N(\mu_D, \sigma_D^2)$ distribution.
Let H_0: $\mu_D = 0$ H_a: $\mu_D \neq 0$..
Reject H_0 for $| T | > t_{0.95} = 1.943$, the 95th quantile of the t-distribution with 6 degrees of freedom. We find $T = \bar{D} / \sqrt{S^2 / 7} = 2.43/1.901 = 1.278$.
p-value = 0.248. Do not reject at $\alpha = 0.10$ level.

b) The ranks of the absolute values, after omitting the zero, are: 5.5, 1, 5.5, 4, 4, 4. The sum of the ranks of the negative values is $W_- = 8$. Then $P(W_- \leq 8) = (1+7+6+3)/2^6 = 17/64$, so the observed p-value is 34/64. Do not reject.

9) a) Let $w_i = x_i + 1$, so $Y_i = \beta w_i + \varepsilon_i$. Let $Q(\beta) = \Sigma (Y_i - \beta w_i)^2$. We want to minimize $Q(\beta)$. Differentiating wrt β and setting $Q'(\beta) = 0$, we get the least squares estimator $\hat{\beta} = \Sigma w_i Y_i / K$, where $K = \Sigma w_i^2$.

b) $E(\hat{\beta}) = \beta \Sigma w_i^2/K = \beta$.   c) $Var(\hat{\beta}) = \Sigma w_i^2 \sigma^2/K^2 = \sigma^2/K$

d) $\hat{\beta} = ½$, $S^2 = 4/9$, T = (1/2-1)/(S/K) = -3, p-value = 2(.0075) = 0.015.

10) a) n = 144, b) n = 144, c) 102
d) The procedure used to determine the interval has the property that 95% of all possible intervals determined when random samples are taken will produce intervals containing the parameter $\mu_X - \mu_Y$. We do not know whether this interval contains the parameter.