Partial solution to Sample Questions for Exam II
ST422 - Introduction to Mathematical Statistics-II

7. Let $Z_1, Z_2 \sim^{i.i.d} N(0,1)$ and $Y_1, Y_2 \sim^{i,i,d} N(1,1)$. Suppose $Z_i's$ are independent of the $Y_j's$. Find the distribution of the following random variables:

   (a) $(Z_1 + Z_2)/\sqrt{[(Y_2 - Y_1)^2 + (Z_2 - Z_1)^2]/2}$.

   (b) $[(Y_1 - Y_2)^2 + (Z_1 - Z_2)^2 + (Z_1 + Z_2)^2]/2$.

   (c) $(Y_2 + Y_1 - 2)^2/(Y_2 - Y_1)^2$.

Solution.

(a) Verify that,
$(Z_1 + Z_2)/\sqrt{2} \sim N(0,1)$.
$(Y_2 - Y_1)/\sqrt{2} \sim N(0,1) \Longrightarrow (Y_2 - Y_1)^2/2 \sim \chi^2_1$.
$(Z_2 - Z_1)/\sqrt{2} \sim N(0,1) \Longrightarrow (Z_2 - Z_1)^2/2 \sim \chi^2_1$.
$Cov(Z_2 + Z_1, Z_2 - Z_1) = 0 \Longrightarrow Z_2 + Z_1$ and $Z_2 - Z_1$ are independent.
(*caution: above fact may NOT be true if Z_1 and Z_2 are not jointly normal.*)
Finally use the fact that,
$(Y_2 - Y_1)^2/2 + (Z_2 - Z_1)^2/2 \sim \chi^2_2$ and is independent of $Z_1 + Z_2$,
to conclude that $(Z_1 + Z_2)/\sqrt{[(Y_2 - Y_1)^2 + (Z_2 - Z_1)^2]/2} \sim t_2$.

(b) Verify that,
$(Z_1 + Z_2)^2/2 \sim \chi^2_1$ and use part (a) to conclude,
$[(Y_1 - Y_2)^2 + (Z_1 - Z_2)^2 + (Z_1 + Z_2)^2]/2 \sim \chi^2_3$.

(c) Verify that, $(Y_2 + Y_1 - 2)/\sqrt{2} \sim N(0,1) \Longrightarrow (Y_2 + Y_1 - 2)^2/2 \sim \chi^2_1$.
$Cov(Y_2 + Y_1, Y_2 - Y_1) = 0 \Longrightarrow Y_2 + Y_1$ and $Y_2 - Y_1$ are independent.
Use part (a) to conclude,
$(Y_2 + Y_1 - 2)^2/(Y_2 - Y_1)^2 \sim F(1,1)$.

8. If $Y_1, Y_2, \ldots, Y_n$ are independently and normally distributed with the same mean but different variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ and assuming that $U = \frac{\sum_{i=1}^n Y_i/\sigma_i^2}{\sum_{j=1}^n (1/\sigma_j^2)}$ and $V = \sum_{i=1}^n (Y_i - U)^2/\sigma_i^2$ are independently distributed, show that U is normal and V has χ_{n-1}^2 distribution.

Solution.

$Y_i \sim N(\mu, \sigma_i^2) \Longrightarrow \frac{Y_i}{\sigma_i^2} \sim N(\frac{\mu}{\sigma_i^2}, \frac{1}{\sigma_i^2})$.
$\Longrightarrow \sum_{i=1}^n \frac{Y_i}{\sigma_i^2} \sim N(\mu \sum_{i=1}^n \frac{1}{\sigma_i^2}, \sum_{i=1}^n \frac{1}{\sigma_i^2})$.
$\Longrightarrow U = \sum_{i=1}^n \frac{Y_i}{\sigma_i^2} / \sum_{i=1}^n \frac{1}{\sigma_i^2} \sim N(\mu, 1/\sum_{i=1}^n \frac{1}{\sigma_i^2})$.

Next notice that, $V = U_2 - U_1$. where,
$U_2 = \sum_{i=1}^n \left(\frac{Y_i - \mu}{\sigma_i}\right)^2 \sim \chi_n^2$ and $U_1 = (U - \mu)^2 \sum_{i=1}^n \frac{1}{\sigma_i^2} \sim \chi_1^2$.

We will use m.g.f. method to find the distribution of V.
Recall that, $m_{U_2} = (1-2t)^{-\frac{n}{2}}$ and $m_{U_1}(t) = (1-2t)^{-\frac{1}{2}}$.

Now assuming that V and U (and hence U_1) independent,
$m_{U_2}(t) = E[e^{tU_2}] = m_V(t) m_{U_1}(t)$
$\Longrightarrow m_V(t) = (1-2t)^{-\frac{n}{2}}/(1-2t)^{-\frac{1}{2}} = (1-2t)^{-\frac{n-1}{2}}$.

Thus, $V \sim \chi_{n-1}^2$.

9. (a) Use the Chebyshev inequality to find how many times a coin must be tossed in order that the probability will be at least 0.90 that $\bar{Y}$ will lie between 0.4 and 0.6?

   (b) How could one determine the number of tosses required in part (a) more accurately, i.e., make the probability very nearly equal to 0.90? What is the number of tosses?

Solution.

(a) Tchebyshev's inequality: $\Pr\left[\left|\frac{\sqrt{n}(\bar{Y}-\mu)}{\sigma}\right| \le k\right] \ge 1 - \frac{1}{k^2}$.
Take $\mu = 0.5, \sigma = \sqrt{0.5(1-0.5)}$. Notice that, $1 - \frac{1}{k^2} = 0.9 \Longrightarrow k = \sqrt{10}$.
Finally, $\Pr[0.4 \le \bar{Y} \le 0.6] = \Pr\left[\left|\frac{\sqrt{n}(\bar{Y}-\mu)}{\sigma}\right| \le k\right] \Longrightarrow \frac{0.1\sqrt{n}}{\sigma} = k$.
Substituting the values of σ, k, we get $n = 250$.

(b) $\Pr[0.4 \le \bar{Y} \le 0.6] \approx \Pr[-\frac{0.1\sqrt{n}}{\sigma} \le Z \le \frac{0.1\sqrt{n}}{\sigma}] = 0.9$,
where $Z \sim N(0,1)$ (using CLT).
Use Table 4, to conclude, $\frac{\sqrt{n}}{5} = 1.645 \Longrightarrow n = 68$.

If you have further questions, see me in my office (209F Patterson Hall).