IE 2027C **Practice Questions for Final Exam** **Spring 2015**

If $X \sim N(0,1)$ then
$$P(X \leq t) \equiv \Phi(t).$$
If $X_i \sim N(\mu_i, \sigma_i^2)$ are mutually independent then $\sum_i X_i \sim N(\sum_i \mu_i, \sum_i \sigma_i^2)$ and $\alpha X_i \sim N(\alpha\mu_i, \alpha^2\sigma_i^2)$. If $Z \sim N(\mu, \sigma^2)$ then
$$\frac{Z-\mu}{\sigma} \sim N(0,1).$$
Therefore
$$P(Z \leq t) = P\left(\frac{Z-\mu}{\sigma} \leq \frac{t-\mu}{\sigma}\right) = \Phi\left(\frac{t-\mu}{\sigma}\right).$$
The argument of Φ is the number of standard deviations (of Z) by which t is bigger than the mean of Z.

Important values of $\Phi(t)$: $\Phi(0) = .5$; $\Phi(1) \approx .84$; $\Phi(1.3) \approx .9$; $\Phi(2) \approx .977$, $\Phi(3) \approx .999$. By symmetry $\Phi(-3) \approx .001$, $\Phi(-2) \approx .023$, etc.

If X_i are i.i.d. each with mean μ and variance σ^2 then define the sample mean as
$$\bar{X} \equiv \sum_{i=1}^{n} \frac{1}{n} X_i.$$
The central limit theorem says that as $n \to \infty$ the distribution of
$$\sqrt{n}\frac{\bar{X}-\mu}{\sigma}$$
converges to the standard Gaussian distribution $N(0,1)$.

1. Random variables X, Y, Z are jointly independently distributed, each Gaussian, with means $3, 7, 11$ respectively and variances $1, 2, 3$ respectively. For example, $Y \sim N(7,2)$. Find the distributions of each of the following.

   (a) $X + Y$
   (b) $X + Y + Z$
   (c) $10Y$
   (d) $2X + 20Y + 100Z$
   (e) $-X$
   (f) $X - Y$
   (g) $6X - 10Y - 20Z$

   Solution:

   (a) $X + Y \sim N(3+7, 1+2) = N(10,3)$
   (b) $X + Y + Z \sim N(3+7+11, 1+2+3) = N(21,6)$
   (c) $10Y \sim N(10 \cdot 3, 10^2 \cdot 1) = N(30,100)$.
   (d) $2X + 20Y + 100Z$. $X \sim N(6,4)$, $20Y \sim N(140,800)$, $100Z \sim N(1100,30000)$. Their sum has distribution $N(6+140+1100, 4+800+30000) = N(1246,30804)$.

(e) $-X \sim N(-1 \cdot 3, (-1)^2 \cdot 1) = N(-3, 1)$.

(f) $X - Y = X + (-Y) \sim N(3 + (-1) \cdot 7, 1 + (-1)^2 \cdot 2) = N(-4, 3)$.

(g) $6X - 10Y - 20Z \sim N(6 \cdot 3 - 10 \cdot 7 - 20 \cdot 11, 36 \cdot 1 + 100 \cdot 2 + 400 \cdot 3) = N(-272, 1436)$.

2. Random variables A, B, C are jointly independently distributed, each Gaussian, with means $-1, 0, 2$ respectively and standard deviations $5, 6, 7$ respectively. What is the distribution of $10A - 3B + 4C$?

   Solution: $10A \sim N(10 \cdot -1, 100 \cdot 5^2) = N(-10, 2500)$. $-3B \sim N(-3 \cdot 0, (-3)^2 \cdot 6^2) = N(0, 324)$. $4C \sim N(4 \cdot 2, 4^2 \cdot 7^2) = N(8, 784)$. *The answer is therefore* $N(-10+0+8, 2500+324+784) = N(-2, 3608)$.

3. Random variables X and Y are independent, each Gaussian distributed. If $2X + 3Y \sim N(10, 50)$ and $X \sim N(-1, 3)$, what are the mean and standard deviation of Y?

   Solution: Let μ, σ_Y *be the mean and standard deviation of* Y*, respectively,. Then* $2 \cdot -1 + 3 \cdot \mu = 10$ *and* $2^2 \cdot 3^2 + 3^2 \cdot \sigma_Y^2 = 50$. *Therefore* $\mu = 4$ *and* $\sigma_Y = \sqrt{14/9} = \sqrt{14}/3$.

4. Random variables W_i are mutually independent each with mean -3 and variance 2. Find the mean, variance, and standard deviation of the sample mean of $W_1, W_2, \ldots, W_n$.

   Solution: Let $\bar{W} \equiv \frac{1}{n}\sum_{i=1}^{n} W_i$ *denote the sample mean. Then* $E[W] = \frac{1}{n}\sum_{i=1}^{n} E[W_i] = \frac{n}{n}(-3) = -3$. *For the variance,* $\sigma_{\bar{W}}^2 = \sigma^2(\frac{1}{n}\sum_{i=1}^{n} \sigma_{W_i}^2 = \frac{1}{n^2} n \cdot 2 = 2/n$. *The standard deviation is therefore* $\sqrt{2/n}$.

5. Random variables W_i are mutually independent each with mean -7 and variance 3. Let $\bar{W}$ denote the sample mean of n of these variables as in the previous problem. What does Chebyshev's inequality tell you about $P(|\bar{W} + 7| \geq 2.5)$?

   Solution: $E[\bar{W}] = -7$. $\sigma_{\bar{W}} = \sqrt{3/n}$. *By Chebyshev's inequality,* $P(|\bar{W} + 7| \geq k\sqrt{3/n}) \leq 1/k^2$. *If* $2.5 = k\sqrt{3/n}$ *then* $k = 2.5\sqrt{n/3}$. *Hence* $P(|\bar{W} + 7| \geq 2.5) \leq 1/k^2 = \frac{3}{6.25n}$.

6. Random variables W_i are mutually independent each with mean 1 and variance 4. Let $\bar{W}$ denote the sample mean of 100 of these variables. What does the central limit theorem tell you about the approximate value of $P(\bar{W} \leq 1.6)$? How about the approximate value of $P(0.6 \leq \bar{W} \leq 1.4)$? How about the approximate values of $P(0.8 \leq \bar{W} \leq 1.4)$ and $P(0.8 \leq \bar{W} \leq 1.2)$?

   Solution: $E[\bar{W} = 1$. $\sigma_{\bar{W}}^2 = 4/100$. $\sigma_{\bar{W}} = .2$. *The event* $\bar{W} \leq 1.6)$ *is the event* $\bar{W} \leq 1 + .6 = E[\bar{W}] + 3\sigma_{\bar{W}}$. *The estimated probability by the CLT is* $\Phi(3) \approx .999$. *The event* $0.6 \leq \bar{W} \leq 1.4$ *is the event* $-2\sigma_{\bar{W}} \leq W - E[W] \leq 2\sigma_{\bar{W}}$. *Its estimated probability is* $\Phi(2) - Phi(-2) \approx .977 - .023 = .954$ *which we usually round down to the commonly used probability value* 0.95. *Similarly,* $P(0.8 \leq \bar{W} \leq 1.4) \approx \Phi(2) - \Phi(-1) \approx .977 - (1 - .84) = 0.82$ *and* $P(0.8 \leq \bar{W} \leq 1.2) \approx .84 - (1 - .84) = .68$.

7. Random variables X_i are mutually independent each with mean 12 and variance $\sigma^2 = 5$. Use Chebyshev's inequality to determine the smallest number of variables n that you can (using only Chebyshev's inequality) needed to make the probability at least $\frac{35}{36}$ that the sample mean differs from 12 by less than 1.6 (in absolute value).

Solution: Let $\bar{X}$ denote the sample mean of variables $X_1, X_2, \ldots, X_n$. Then $E[\bar{X}] = 12$ and $\sigma^2_{\bar{X}} = \frac{1}{n^2}5n = 5/n$.

The complement of the event $|\bar{X} - 12| < 1.6$ is the event $|\bar{X} - 12| \geq 1.6$. Therefore $P(|\bar{X} - 12| < 1.6) = 1 - P(|\bar{X} - 12| \geq 1.6)$.

Hence $P(|\bar{X} - 12| < 1.6) \geq 35/36$ iff $P(|\bar{X} - 12| \geq 1.6) \leq 1 - 35/36 = 1/36$.

By Chebyshev's inequality, $P(|\bar{X} - 12| \geq k\sigma_{\bar{X}}) \leq \frac{1}{k^2}$. We want a value of n such that $\frac{1}{k^2} = \frac{1}{36}$ which is equivalent to $k = 6$. So we want a value of n so that $1.6 = k\sigma_{\bar{X}} = 6\sigma_{\bar{X}}$ or equivalently $\sigma_{\bar{X}} = \frac{1.6}{6} = \frac{4}{15}$. We know that $\sigma_{\bar{X}} = \sqrt{5/n}$. Hence $\sqrt{5/n} = 4/15$ and $5/n = 16/225$ and $n = 225 \cdot 5/16 = 70.3125$. But since n has to be an integer we have to round up to $n = 71$.

8. Random variables X_i are mutually independent each with mean 12 and variance $\sigma^2 = 5$. Use the central limit theorem (CLT) to determine the number of variables n needed to make the probability approximately 0.002 that the sample mean differs from 12 by less than 0.6 (in absolute value). How large must n be to make the probability approximately 95%?

   Solution: The variance of $\bar{X}$ is $5/n$ and the standard deviation is $\sqrt{5/n}$. Let $Z = \frac{\bar{X}-12}{\sqrt{5/n}} \sim N(0,1)$ approximately by the CLT. $P(-3 \leq Z \leq 3) \approx .999 - (1 - .999) = .002$. $|Z| \leq 3$ iff $|\bar{X} - 12| \leq 3\sqrt{5/n}$. Set $0.6 = 3\sqrt{5/n}$ to get $.2^2 = \frac{5}{n}$ or $n = 5 \cdot 225/16 = 125$. To get probability $\approx .95$ we need $-2 \leq Z \leq 2$. Set $0.6 = 2\sqrt{5/n}$ to get $.3^2 = \frac{5}{n}$ or $n \approx 55.56$ which we round to the nearest integer 56 (since we were to get a probability of approximately .95).