Exercises: 8.4-8.5

8.18 Prove that if $X_1, \ldots, X_n$ are independent and each is χ^2 with $\nu_1, \nu_2, \ldots, \nu_n$ degrees of freedom (respectively), then

$$Y = \sum_{i=1}^{n} X_i^2$$

is χ^2 with $\nu_1 + \nu_2 + \ldots + \nu_n$ degrees of freedom.

The proof goes quickly with the Moment Generating Function technique (can be used because the X_i are all independent).

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} (1-2t)^{-\frac{1}{2}\nu_i} = (1-2t)^{-\frac{1}{2}(\nu_1+\nu_2+\ldots+\nu_n)}$$

8.19 If X_1 is χ^2 with ν_1 degrees of freedom, and $X_1 + X_2$ is χ^2 with $\nu > nu_1$ degrees of freedom, then X_2 is χ^2 with $\nu - \nu_1$ degrees of freedom:

$$M_{X_1}(t)M_{X_2}(t) = M_{X_1+X_2}(t) \qquad \Rightarrow \qquad (1-2t)^{-\frac{1}{2}\nu_1} M_{X_2}(t) = (1-2t)^{-\frac{1}{2}\nu} \quad \Rightarrow$$

$$M_{X_2}(t) = (1-2t)^{-\frac{1}{2}(\nu-\nu_1)}$$

Therefore, X_2 is χ^2 with $\nu - \nu_1$ degrees of freedom.

8.20 Verify the identity:

$$\sum_{i=1}^{n} (X_i - \mu)^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2$$

Start with the left side of the equation and add/subtract $\bar{X}$, then expand the result and simplify:

$$\sum_{i=1}^{n} (X_i - \mu)^2 = \sum_{i=1}^{n} \left[(X_i - \bar{X}) + \left(\bar{X} - \mu\right)\right]^2 =$$

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 + \sum_{i=1}^{n} \left(\bar{X} - \mu\right)^2 + 2\sum_{i=1}^{n} (X_i - \bar{X})(\bar{X} - \mu) =$$

The $(\bar{X} - \mu)$ term does not depend on i, and so can be factored out of the sum:

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 + n\left(\bar{X} - \mu\right)^2 + 2(\bar{X} - \mu)\sum_{i=1}^{n} (X_i - \bar{X})$$

So, it all boils down to: Show that $\sum_{i=1}^{n} (X_i - \bar{X}) = 0$, and the desired result will follow.

$$\sum_{i=1}^{n} (X_i - \bar{X}) = \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \bar{X} = n \cdot \frac{1}{n} \sum_{i=1}^{n} X_i - \bar{X} \sum_{i=1}^{n} 1 = n\bar{X} - n\bar{X} = 0$$

8.22 The idea is to invoke the Central Limit Theorem, and so we want to show that:

$$\frac{Y_n}{n} = \bar{X} \qquad \mu = 1 \qquad \sigma = \sqrt{2}$$

Given the setup of the problem,

$$Y_n = X_1 + X_2 + \ldots + X_n$$

where X_i are iid with χ^2 and $\nu = 1$. The mean and variance of each distribution associated with X_i is then $\mu = 1$ and $\sigma^2 = 2$.

Therefore,

$$\frac{Y_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

and the mean of $\frac{Y_n}{n} = \mu = 1$ and the std of Y_n/n is $\sqrt{2}$, as desired. The conclusion is the CLT.

8.23 So we don't get the ν's mixed up, I'm changing the notation of the problem to read that X has N degrees of freedom, and

$$\frac{X - N}{\sqrt{2N}} \rightarrow N(0, 1)$$

We recall that if $X_1, \ldots, X_N$ are iid - χ^2 with $\nu = 1$, then we could express X, a χ^2 distribution with N degrees of freedom, as:

$$X = X_1 + X_2 + \ldots + X_N$$

So using the notation of the previous problem,

$$X = Y_N \quad \Rightarrow \quad \frac{X}{N} = \frac{Y_N}{N}$$

Now, simply use some algebra (multiply numerator and denominator by N):

$$\frac{\frac{Y_N}{N} - 1}{\sqrt{2/N}} = \frac{\frac{X}{N} - 1}{\sqrt{2/N}} = \frac{X - N}{\sqrt{2N}}$$

Therefore, by 8.23, the result follows.

8.24 Find $P(X > 68.0)$ if X is χ^2 with 50 degrees of freedom, using the approximation in 8.24:

$$X > 68 \quad \Rightarrow \quad \frac{X - 50}{\sqrt{100}} > 1.8$$

so

$$P(X > 68.0) \approx P(Z > 1.8)$$

where Z comes from $N(0, 1)$. Using the table,

$$\frac{1}{2} - P(0 < Z < 1.8) = \frac{1}{2} - 0.4641 = 0.0359$$

8.64 $n = 81$, $\mu = 128$, $\sigma = 6.3$. Estimate the probability that $\bar{X}$ is NOT between 126.6 and 129.4.

(a) With Chebyshev,

$$126.6 < \bar{X} < 129.4 \quad \Rightarrow \quad -1.4 < \bar{X} - \mu < 1.4 \quad \Rightarrow \quad |\bar{X} - \mu| < 1.4$$

and

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

so we need to find k. We can set up an equation to find k (Recall that the standard deviation of $\bar{X}$ is $\sigma/\sqrt{n}$, and not σ):

$$1.4 = k \cdot \frac{6.3}{\sqrt{81}} \quad \Rightarrow \quad k = 2$$

so that the desired estimate is $\frac{1}{4}$.

(b) Using the normal distribution for the approximation, we let Z be $N(0, 1)$, and:

$$P\left(\left|\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right| < 2\right) \approx P(|Z| < 2) = 2P(0 < Z < 2) = 2 \cdot 0.4772 = 0.9544$$

So the probability of NOT being in that interval is $1 - 0.9544 = 0.0456$.

8.77 The claim that the variance of a normal population is $\sigma^2 = 4$ is to be rejected if the variance of a random sample of size 9 exceeds 7.7535. What is the probability that this claim will be rejected even thought $\sigma^2 = 4$?

Consider the distribution of

$$Y = \frac{(n-1)S^2}{\sigma^2}$$

We know that this is χ^2 with 8 degrees of freedom. NOTE: The random variable is S^2, everything else is constant...

So, we ask for:

$$P(S^2 > 7.7535) = P\left(\frac{8 \cdot S^2}{4} > 2 \cdot 7.7535\right) = P(Y > 15.507)$$

In the χ^2 table, in the row with 8 degrees of freedom, we see that $\alpha = 0.05$. Therefore, the probability is 0.05, or 0.5%.

8.78 $n = 25$ from normal, $\bar{x} = 47$, $s = 7$. Use the $t-$distribution to see if it is reasonable to claim that the population mean is $\mu = 42$.

$$t_{\alpha,24} = \frac{47 - 42}{7/5} = 3.57$$

In the table (in the row with $\nu = 24$) we see that the probability that $P(X > 2.797) = 0.005$... Therefore, it is very unlikely that the population mean is 42.

Added (1) A tensile strength test was performed in order to determine the strength of a particular adhesive. The data obtained was: $\{16, 14, 19, 18, 19, 20, 15, 18, 17, 18\}$.

(a) Compute the sample mean and variance.

$$\bar{X} = \frac{1}{10}(16 + 14 + \ldots + 18) = 17.4$$

$$S^2 = \frac{1}{9}\left((-1.4)^2 + (-3.4)^2 + \ldots + (0.6)^2\right) = 3.6$$

(b) Assuming these came from a normal distribution, find an interval in which the true mean would be included 90% of the time.

By "these", we meant that the random sample was iid, coming from a normal distribution (so we can use the $t-$distribution). We want an interval for the population mean μ:

$$\left|\frac{\mu - \bar{X}}{s/\sqrt{n}}\right| < t_{\alpha,\nu} \quad \Rightarrow \quad \bar{X} - t_{\alpha,\nu} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha,\nu}\frac{s}{\sqrt{n}}$$

In the notation from class, $p = 0.9$, so $\alpha = \frac{1-p}{2} = 0.05$, $\nu = 9$. This gives $t_{\alpha,\nu}$ from the table as: 1.833

$$17.4 - 1.833 \cdot \sqrt{\frac{3.6}{10}} < \mu < 17.4 + 1.833 \cdot \sqrt{\frac{3.6}{10}} \quad \Rightarrow \quad \mu \in 17.4 \pm 1.0998$$

Added (2) Prior to the 1999-2000 NBA season, rules were changed to increasing scoring and make the games more exciting. Previously, the league averaged 183.2 points per game. If we sampled 25 games afterward and the average was $\bar{x} = 195.88$ and $s = 20.27$, should we conclude that they succeeded (at least in the increased scoring)?

This setup is a lot like 8.78, except for the conclusion. Here we see that:

$$\frac{195.88 - 183.2}{20.27/5} \approx 3.12$$

In the table (with 24 degrees of freedom), we see that the probability is very small (negligible) that we would have gotten a sample mean of 195.5.88 if the population mean was still 183.2, therefore we would conclude that there is a strong probability that the population mean has increased.