Homework 2 Solutions

1. [§5-26] Suppose that a basketball player can score on a particular shot with probability .3. Use the central limit theorem to find the approximate distribution of $S$, the number of successes out of 25 independent shots. Find the approximate probabilities that $S$ is less than or equal to 5, 7, 9, and 11 and compare these to the exact probabilities. \textit{(Hint: Let $X_1, X_2, \ldots, X_{25}$ be the indicator random variables of the 25 shots, that is, $X_i = 1$ if the player scores on the $i$th shot and $X_i = 0$ otherwise.)}

Let $X_1, X_2, \ldots, X_{25}$ be the indicator random variables of the 25 shots, that is, $X_i = 1$ if the player scores on the $i$th shot and $X_i = 0$ otherwise. Then $X_i$’s are i.i.d. Bernoulli random variables with parameter $p = 0.3$. It follows that the common mean and variance is

$$
\mu = \mathbb{E}[X_1] = p = 0.3 \quad \text{and} \quad \sigma^2 = \text{Var}[X_1] = p(1 - p) = 0.21.
$$

Since $S = X_1 + X_2 + \cdots + X_{25}$, the central limit theorem tells us that $S$ is approximately normal with mean $n\mu = 25(0.3) = 7.5$ and variance $n\sigma^2 = 25(0.21) = 5.25$, that is

$$
S \sim N(7.5, 5.25), \quad \text{approximately.}
$$

Let $Z \sim N(0, 1)$. Then, we have

$$
P(S \leq 5) = \mathbb{P}\left( \frac{S - n\mu}{\sigma \sqrt{n}} \leq \frac{5 - n\mu}{\sigma \sqrt{n}} \right) = \mathbb{P}\left( \frac{S - 7.5}{\sqrt{5.25}} \leq \frac{5 - 7.5}{\sqrt{5.25}} \right) \\
\approx \mathbb{P}(Z \leq -1.09) = 0.1379;
$$

$$
P(S \leq 7) = \mathbb{P}\left( \frac{S - n\mu}{\sigma \sqrt{n}} \leq \frac{7 - n\mu}{\sigma \sqrt{n}} \right) = \mathbb{P}\left( \frac{S - 7.5}{\sqrt{5.25}} \leq \frac{7 - 7.5}{\sqrt{5.25}} \right) \\
\approx \mathbb{P}(Z \leq -0.22) = 0.4129;
$$

$$
P(S \leq 9) = \mathbb{P}\left( \frac{S - n\mu}{\sigma \sqrt{n}} \leq \frac{9 - n\mu}{\sigma \sqrt{n}} \right) = \mathbb{P}\left( \frac{S - 7.5}{\sqrt{5.25}} \leq \frac{9 - 7.5}{\sqrt{5.25}} \right) \\
\approx \mathbb{P}(Z \leq 0.65) = 0.7422; \quad \text{and}
$$

$$
P(S \leq 11) = \mathbb{P}\left( \frac{S - n\mu}{\sigma \sqrt{n}} \leq \frac{11 - n\mu}{\sigma \sqrt{n}} \right) = \mathbb{P}\left( \frac{S - 7.5}{\sqrt{5.25}} \leq \frac{11 - 7.5}{\sqrt{5.25}} \right) \\
\approx \mathbb{P}(Z \leq 1.53) = 0.9370.
$$

On the other hand, we have $S \sim Bin(25, 0.3)$, that is, $S$ is a binomial random variable with parameters $n = 25$ and $p = 0.3$. According the binomial cdf table, we have

$$
P(S \leq 5) = 0.193, \quad P(S \leq 7) = 0.512, \quad P(S \leq 9) = 0.811, \quad \text{and} \quad P(S \leq 11) = 0.956.
$$
Some remark: note that we have an under estimation here when using normal to approximate binomial. Recall that it is suggested to use the "continuity adjustment" when approximating binomial by normal if $n$ is not large enough:

$$P(S \leq x) \approx P \left( Z \leq \frac{x - n\mu + 0.5}{\sigma\sqrt{n}} \right).$$

For example,

$$P(S \leq 5) \approx P \left( Z \leq \frac{5 - 7.5 + 0.5}{\sqrt{5.25}} \right) = 0.1914.$$

2. Assume $X \sim N(0, 1)$.

(a) Find $P(X > 1.5)$;
(b) Find the value $c$ such that $P(X < c) = 0.05$; and
(c) Find the value $d$ such that $P(|X| < d) = 0.9$.

(a) $P(X > 1.5) = 1 - P(X \leq 1.5) = 1 - 0.9332 = 0.0668$.

R function:

```
1 - pnorm(1.5)
```

Remark: the syntax for the R function of the cdf of normal distributions is as follows:

```
pnorm(x, mu, sigma),
```

and this gives $P(X \leq x)$ for $X \sim N(\mu, \sigma^2)$. Here mu= $\mu$ and sigma= $\sigma$. When these two parameters are omitted, R assumes $\mu = 0$ and $\sigma = 1$.

(b) $c = -1.645$.

R function:

```
qnorm(0.05)
```

Remark: the syntax for the R function of the quantile (inverse cdf) of normal distributions is as follows:

```
qnorm(p, mu, sigma),
```
and this gives the value \( c \) such that \( P(X \leq c) = p \) for \( X \sim N(\mu, \sigma^2) \).

(c) By symmetry we have

\[
0.9 = P(|X| < d) = P(-d < X < d) = P(X < d) - P(X < -d) \\
= P(X < d) - (1 - P(X < d)) \\
= 2P(X < d) - 1.
\]

Thus \( P(X < d) = 0.95 \) and \( d = 1.645 \).

3. If \( T \sim t_{20} \),

(a) find \( P(T < -1.725) \);
(b) find \( \tau_1 \) such that \( P(T > \tau_1) = .05 \); and
(c) find \( \tau_2 \) such that \( P(|T| < \tau_2) = .9 \).

(a) \( P(T < -1.725) = 0.05 \).

R function:

\[
\text{pt}(-1.725, 20)
\]

Remark: the syntax for the R function of the cdf of \( t \)-distribution with degrees of freedom \( n \) is as follows:

\[
\text{pt}(x, df)
\]

and this gives \( P(X \leq x) \) for \( X \sim t_n \).

(b) Since \( P(T > \tau_1) = .05 \), we have \( P(T < \tau_1) = 0.95 \). Thus \( \tau_1 = 1.725 \).

R function:

\[
\text{qt}(0.95, 20)
\]

Remark: the syntax for the R function of the quantile (inverse cdf) of \( t \)-distribution with degrees of freedom \( n \) is as follows:

\[
\text{qt}(p, df)
\]

and this gives the value \( c \) such that \( P(X \leq c) = p \) for \( X \sim t_n \).
(c) the $t$ distribution is symmetric about 0, we have

$$0.9 = \Pr(|T| < \tau_2) = \Pr(T < \tau_2) - \Pr(T < -\tau_2) = 2\Pr(T < \tau_2) - 1.$$ 

Thus $\Pr(T < \tau_2) = 0.95$ and $\tau_2 = 1.725$.

4. If $U \sim \chi^2_6$,

(a) find $\Pr(U \leq 12.59)$; and

(b) the value $c$ such that $\Pr(U > c) = 0.025$.

(a) $\Pr(U \leq 12.59) = 0.95.$

R function:

```
pchisq(12.59,6)
```

Remark: the syntax for the R function of the cdf of $\chi^2$-distribution with degrees of freedom $n$ is as follows:

```
pchisq(x,df)
```

and this gives $\Pr(X \leq x)$ for $X \sim \chi^2_n$.

(b) Since $\Pr(U > c) = .025$, we have $\Pr(U \leq c) = 0.975$. Thus $c = 14.45$.

R function:

```
qchisq(0.975,6)
```

Remark: the syntax for the R function of the quantile (inverse cdf) of $\chi^2$-distribution with degrees of freedom $n$ is as follows:

```
qt(p,df)
```

and this gives the value $c$ such that $\Pr(X \leq c) = p$ for $X \sim \chi^2_n$.

5. If $W \sim F_{2,3}$,

(a) find $\Pr(W \geq 9.55)$; and

(b) the value $c$ such that $\Pr(W > c) = 0.025$. 

4
(a) \( \mathbb{P}(W \geq 9.55) = 1 - \mathbb{P}(W < 9.55) = 1 - 0.95 = 0.05. \)

R function:
\[
\text{pf}(9.55, 2, 3)
\]

Remark: the syntax for the R function of the cdf of \( \mathcal{F} \)-distribution with degrees of freedom \( n \) (numerator) and \( m \) (denominator) is as follows:
\[
\text{pf}(x, n, m)
\]
and this gives \( \mathbb{P}(X \leq x) \) for \( X \sim \mathcal{F}_{n,m} \).

(b) Since \( \mathbb{P}(W > c) = .025 \), we have \( \mathbb{P}(W \leq c) = 0.975 \). Thus \( c = 16.04 \).

R function:
\[
\text{qf}(0.975, 2, 3)
\]

Remark: the syntax for the R function of the quantile (inverse cdf) of \( \mathcal{F} \)-distribution with degrees of freedom \( n \) (numerator) and \( m \) (denominator) is as follows:
\[
\text{qf}(p, n, m)
\]
and this gives the value \( c \) such that \( \mathbb{P}(X \leq c) = p \) for \( X \sim \mathcal{F}_{n,m} \).

6. [6-3] Let \( \overline{X} \) be the average of a sample of 16 independent normal random variables with mean 0 and variance 1. Determine \( c \) such that
\[
\mathbb{P}(|\overline{X}| < c) = .5.
\]

(Review: if the pdf of a random variable \( Y \) is symmetric (about its mean \( \mu \)), then \( \mathbb{P}(Y - \mu < a) = \mathbb{P}(Y - \mu > -a) \)).
First, we have $X \sim N(0, 1/16)$. Thus

$$0.5 = \Pr(|X| < c) = \Pr(X < c) - \Pr(X < -c)$$

$$= \Pr(X < c) - (1 - \Pr(X < c)) = 2\Pr(X < c) - 1.$$ 

This means that we need to find $c$ such that $\Pr(X < c) = 0.75$. Now

$$0.75 = \Pr(X < c) = \Pr\left(\frac{X - 0}{\sqrt{1/16}} < \frac{c - 0}{\sqrt{1/16}}\right) = \Pr(Z < 4c)$$

implies that $4c = 0.675$, or equivalently, $c = 0.169$.

7. [§6-5] Show that if $X \sim F_{n,m}$, then $X^{-1} \sim F_{m,n}$.

Since $X \sim F_{n,m}$, we have

$$X \overset{d}{=} \frac{U/n}{V/m},$$

where $U$ and $V$ are independent random variables with $U \sim \chi^2_n$ and $V \sim \chi^2_m$, and $\overset{d}{=} \text{ means equal in distribution.}$ It follows immediately that

$$X^{-1} \overset{d}{=} \frac{V/m}{U/n}.$$

Since $\frac{V/m}{U/n} \sim F_{m,n}$ by definition, we have $X^{-1} \sim F_{m,n}$.
8. Let $X_1, X_2, \ldots, X_{11}$ be i.i.d. random variables with $X_i \sim N(\mu, \sigma)$ for $1 \leq i \leq 11$, and $S^2$ be the sample variance. Use the chi-square distribution to calculate $P(0.394 < S^2/\sigma^2 < 1.831)$.

Since $\frac{(n - 1)S^2}{\sigma^2} \sim \chi_{n-1}^2$, we have

$$P(0.394 < S^2/\sigma^2 < 1.831) = P \left( 0.394 \cdot (11 - 1) < \frac{(11 - 1)S^2}{\sigma^2} < 1.831 \cdot (11 - 1) \right)$$

$$= P(3.94 < X < 18.31)$$

$$= P(X < 18.31) - P(X \leq 3.94)$$

$$= 0.95 - 0.05$$

$$= 0.9,$$

where $X \sim \chi_{10}^2$. 
