Solutions to Homework 7

1. (a) **Suppose we toss a fair ($p = 0.5$) coin n times. Use the Chebychev inequality to find how many times a coin must be tossed in order that the probability will be at least 0.90 that the observed frequency of heads will lie between 0.4 and 0.6.**

Answer: Define, $X_i \sim Bernoulli(p)$. The observed frequency of heads is $\bar{X}_n = \sum_{i=1}^n X_i/n$, where $\sum_{i=1}^n X_i \sim Bin(n,p)$ and $p = 0.5$. Then,

$$E[\bar{X}_n] = \mu = 0.5,\ Var(\bar{X}_n) = \frac{\sigma^2}{n} = \frac{p(1-p)}{n} = \frac{1}{4n}.$$

Chebyshev's Inequality gives

$$P(|\bar{X}_n - \mu| \geq \epsilon) \leq \frac{Var(\bar{X}_n)}{\epsilon^2}.$$

Plugging in the values from the problem description, we want to solve for n in,

$$P(|\bar{X}_n - 0.5| \geq 0.1) \leq \frac{1}{4n0.1^2} \leq 0.1,$$

which gives us a bound of $n \geq 250$.

(b) **Alternatively, use the normal approximation to determine the number of of tosses required in part a). How do the results in part a) and b) compare?**

Answer: By the CLT, $\bar{X}_n \sim N(\mu, \sigma^2)$, so

$$P(|\bar{X}_n - \mu| \geq \epsilon) \approx 2(1 - \Phi(\sqrt{n}\epsilon/\sigma))$$

Plugging in the values from the problem description, we want to solve for n,

$$P(|\bar{X}_n - 0.5| \geq 0.1) \approx 2(1 - \Phi(\sqrt{n}0.1/\frac{1}{2})) = 0.1$$

$$n \approx (\frac{\Phi^{-1}(1-0.05)}{0.2})^2 \approx 68$$

This is a more precise bound.

2. **Let sample space $S = [0,1]$ with the uniform density function and define $Z(s) = s$ and $Z_1, Z_2 \ldots$ as**

$$Z_1(s) = s + I_{[0,1]}(s),\quad Z_2(s) = s + I_{[0,1/2]}(s),\quad Z_3(s) = s + I_{[1/2,1]}(s)$$
$$Z_4(s) = s + I_{[0,1/3]}(s),\quad Z_5(s) = s + I_{[1/3,2/3]}(s),\quad Z_6(s) = s + I_{[2/3,1]}(s)$$

(a) **Does $Z_n \to^{\mathcal{P}} Z$?**

Yes, Because

$$P(|Z_n - Z| \geq \epsilon) = P(I_D(s) \geq \epsilon),$$

so we have an interval D of s values that goes to 0.

(b) **Does $Z_n \rightarrow^{a.s.} Z$?**

No, there is no value of $s \in S$, s.t. $Z_n(s) \rightarrow s = Z(s)$. For every s $Z_n(s)$ alternates between the values s and $s+1$. Example, if $s = 3/8$, $Z_1(s) = 1\ 3/8$, $Z_2(s) = 1\ 3/8$, $Z_3(s) = 3/8$, $Z_4(s) = 3/8$, $Z_5(s) = 1\ 3/8$, $Z_6(s) = 3/8$. No pointwise convergence.

3. Let F be a cdf and let F^{-1} denote its left-continuous inverse.

(a) Show that $F \circ F^{-1}(t) \geq t$ for all $0 < t < 1$ with equality iff t is in the range of F. Hence $P(F(X) \leq t) \leq t$.

Answer: Let $x_0 = inf\{x : F(x) \geq t\} = F^{-1}(t)$. Now $F(x_0) \geq t$ by right continuity of F; so:

$$F(F^{-1}(t)) = F(x_0) \geq t$$

- Now show $F \circ F^{-1}(t) = t \Leftrightarrow t \in R$ where R is the range of F.

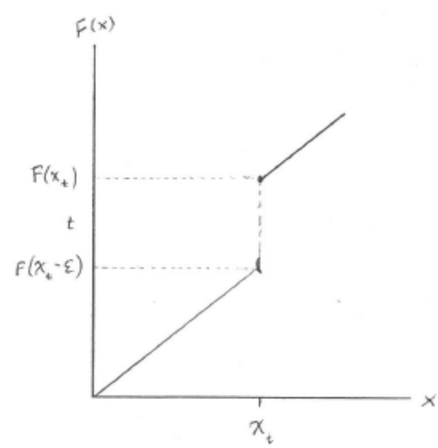
$\Rightarrow$ Let $x_0 = F^{-1}(t)$, then suppose $F(F^{-1}(t)) = t$ which implies $F(x_0) = t \Rightarrow t \in R$.

$\Leftarrow$ Let $t \in R$ and let $T = \{x : F(x) \geq t\}$. Since $t \in R$ there exists x_t such that $F(x_t) = t$ so that $X_t \in T$. Then for any other $x \in T$, it follows that $F(x) \geq F(x_t) = t$. If $x_0 = \inf T$, then $F(x_0) \geq t$ by right continuity of F. So by properties of infimum, $F(x_0) = t$ since $F(x_t) = t$. It follows that $F(F^{-1}(t)) = F(x_0) = t$.

- Show that the above implies $P(F(x) \leq t) \leq t$. Let R denote the range of F.
  - Suppose $t \in R$ then from above: $F(F^{-1}(t)) = t$ and $F(F^{-1}(t)) = P(x \leq F^{-1}(t))$. Since F is non-decreasing, it can be applied to both sides of the inequality:

$$P(x \leq F^{-1}(t)) = P(F(x) \leq F(F^{-1}(t)) = P(F(x) \leq t)$$

    Thus for $t \in R$, $P(F(x) \leq t) = t$.
  - Now suppose $t \notin R$.



By the above picture (with $x_t = F^{-1}(t)$):

$$P(F(x) \leq t) = P(F(x) < t) = \lim_{\epsilon \to 0} P(F(x) \leq F(x_t - \epsilon))$$

Now, $\lim_{\epsilon \to 0} F(x_t - \epsilon) < t$ which implies $P(F(x) \leq t) < t$ for $t \notin R$.

  - Hence, for all t, $P(F(x) \leq t) \leq t$.

(b) Show that $F^{-1} \circ F(x) \leq x$ for all $x \in (-\infty, \infty)$ with strick inequality iff $F(x - \epsilon) = F(x)$ for some $\epsilon > 0$. Hence $P(F^{-1} \circ F(X) \neq X) = 0$ for $X \sim F$.

3b(i) First note the following:

From the definition: $F^{-1}(F(x_0)) = inf\{x : F(x) = F(x_0)\}$
Let $A_0 = \{x : F(x) = F(x_0)\}$, and let $x_0^* = inf A_0$.
It follows that $x_0^* \leq x_0$; hence, $x_0^* = F^{-1}(F(x_0)) \leq x_0$.
Since F is defined for all $x \in (-\infty, \infty)$, and since $0 \leq F(x) \leq 1 \; \forall x$, it follows that $A_0 \neq \emptyset$ for all $x_0 \in (-\infty, \infty)$, and therefore holds for all $x \in (-\infty, \infty)$.

Now show that $F^{-1} \circ F(x) < x \Leftrightarrow F(x - \epsilon) = F(x)$ for some $\epsilon > 0$.

$\Rightarrow$ If $F^{-1}(F(x_0)) < x_0$, then $x_0^* = inf\{x : F(x) = F(x_0)\} < x_0$. Thus, there exists $\epsilon > 0$ such that $x_0^* \leq x - \epsilon < x$. By right-continuity of F, $F(x_0^*) = F(x)$, and since F is non-decreasing $F(x_0^*) = F(x - \epsilon) = F(x)$.

$\Leftarrow$ Suppose $F(x_0 - \epsilon) = F(x_0)$ for some $\epsilon > 0$.
Then $F^{-1}(F(x_0)) = x_0^* = inf\{x : F(x) = F(x_0)\} = inf A_0$ which implies $x_0^* \leq x$ for all $x \in A_0$.
So $F(x_0 - \epsilon) = F(x_0) \Rightarrow (x_0 - \epsilon) \in A_0$, and $x_0 \in A_0$; thus, $x_0^* \leq x_0 - \epsilon < x_0$ and therefore $x_0^* = F^{-1}(F(x_0)) < x_0$.

3b(ii) Show that the results of 3.b(i) imply $P(F^{-1}(F(x)) \neq x) = 0$ for $X \sim F$.

From 3b(i): $F^{-1}(F(x)) \neq x$ implies $F^{-1}(F(x)) < x$, and $F^{-1}(F(x)) < x \Leftrightarrow F(x - \epsilon) = F(x)$ for some $\epsilon > 0$.

Let $A = \{x : F^{-1}(F(x)) < x\}$ and let $B^* = \{F(x) : x \in A\}$.

Note $A = \{x : F(x - \epsilon) = F(x)$ for some $\epsilon > 0\}$ (i.e., the flat parts of the cdf).

There are at most a countable number of different values in B^* because there can be at most a countable number of flat regions in the cdf.

Let $B = \{F(x_1), F(x_1), ..., F(x_k), ...\}$, and let $A_i = \{x : F(x) = F(x_i)$; then $\bigcup_{i=1}^{\infty} A_i = A$.

Let $m_i = inf A_i$ and $M_i = sup A_i$, because these regions are flat $F(m_i) = F(M_i)$; so $P(A_i) = F(M_i) - F(m_i) = 0$, and $P(\bigcup_{i=1}^{\infty} A_i) = 0$ so $P(F^{-1}(F(x)) \neq x) = 0$.

4. BD 5.3.13, pg 351*

(a) **Show that if n is large $\sqrt{S_n} - \sqrt{n}$ has approximately a $N(0, \frac{1}{2})$ distribution. This is known as <u>Fisher's approximation</u>.**
Since S_n has a χ^2_n distribution, $S_n = \sum_{i=1}^{n} X_i^2$, where $X_i \sim N(0,1)$ and $E[X_i^2] = \mu = 1$ and $Var(X_i^2) = \sigma^2 = 2$.
To use the Delta Method, define $\bar{X}_n = S_n/n$ and $h(t) = \sqrt{t}$, which is continuous and $h'(t) = \frac{1}{2}t^{-1/2}$. Then,

$$\sqrt{n}(h(\bar{X}_n) - h(\mu)) \to^{\mathcal{L}} N(0, \sigma^2[h^{(1)}(\mu)]^2)$$

$$\sqrt{n}(\sqrt{S_n/n} - \sqrt{\mu}) \to^{\mathcal{L}} N(0, 2[\frac{1}{2}\mu^{-1/2}]^2)$$

Since $\mu = 1$,

$$\sqrt{S_n} - \sqrt{n} \to^{\mathcal{L}} N(0, \frac{1}{2}).$$

(b) **From a) deduce the approximation $P[S_n \le x] \approx \Phi(\sqrt{2x} - \sqrt{2n})$.**

$$P[S_n \le x] = P[\sqrt{S_n} \le \sqrt{x}] = P[\sqrt{S_n} - \sqrt{n} \le \sqrt{x} - \sqrt{n}]$$
$$= P[\sqrt{2}(\sqrt{S_n} - \sqrt{n}) \le \sqrt{2}(\sqrt{x} - \sqrt{n})]$$

Since, $\sqrt{2}(\sqrt{S_n} - \sqrt{n})$ is approximately $N(0,1)$, then

$$P[S_n \le x] \approx \Phi(\sqrt{2x} - \sqrt{2n}).$$

(c) **Compare the approximation of b) with the central limit theorem approximation $P[S_n \le x] \approx \Phi((x-n)/\sqrt{2n})$ and the exact values of $P[S_n \le x]$ from the χ^2 table for $x = x_{0.90}$, $x = x_{0.99}$, $n = 5, 10, 25$. Here x_q denotes the qth quantile of the χ^2_n distribution.**

For $n = 5$, $x_{0.90} = 9.236357$ and $x_{0.99} = 15.08627$. Approximate and exact values, for $P[S_n \le x]$:

	Exact (χ^2_n)	Fisher's Approx ($\sqrt{S_n}$)	CLT (S_n)
$x_{0.90}$	0.90	0.8719614	0.909821
$x_{0.99}$	0.99	0.9901148	0.9992876

For $n = 10$, $x_{0.90} = 15.98718$ and $x_{0.99} = 23.20925$. Approximate and exact values, for $P[S_n \le x]$:

	Exact (χ^2_n)	Fisher's Approx ($\sqrt{S_n}$)	CLT (S_n)
$x_{0.90}$	0.90	0.8814867	0.9096779
$x_{0.99}$	0.99	0.9903833	0.99843

For $n = 25$, $x_{0.90} = 34.38159$ and $x_{0.99} = 44.31410$. Approximate and exact values, for $P[S_n \le x]$:

	Exact (χ^2_n)	Fisher's Approx ($\sqrt{S_n}$)	CLT (S_n)
$x_{0.90}$	0.90	0.8890116	0.9077054
$x_{0.99}$	0.99	0.9904401	0.9968470

Compared to the CLT approximation, Fisher's Approximation is a better approximation to the probabilities for $x_{0.99}$. In contrast, the CLT approximation better approximates the probabilities for $x_{0.90}$. Both approximations improve as n increases.

5. (BD 5.3.28, pg 354) **Suppose $X_1, \ldots, X_{n_1}$ and $Y_1, \ldots, Y_{n_2}$ are as in Section 4.9.3 independent samples with $\mu_1 = E(X_1)$, $\sigma_1^2 = Var(X_1)$, $\mu_2 = E(X_2)$, $\sigma_2^2 = Var(X_2)$. We want to study the behavior of the two-sample pivot $T(\Delta)$ of Example 4.9.3, if $n_1, n_2 \to \infty$, so that $n_1/n \to \lambda$, $0 < \lambda < 1$.**

Show that $P[T(\Delta) \le t] \to \Phi(t[(\lambda\sigma_1^2 + (1-\lambda)\sigma_2^2)/((1-\lambda)\sigma_1^2 + \lambda\sigma_2^2)]^{\frac{1}{2}})$.

The two-sample t-statistic for testing $H_0 : \mu_1 = \mu_2$ ($\Delta = 0$), versus $H_1 : \mu_2 > \mu_1$ is

$$T(\Delta) = S_n = \sqrt{\frac{n_1 n_2}{n}} \frac{\bar{Y} - \bar{X} - \Delta}{s},$$

where $s^2 = \frac{1}{n-2}(\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2$ and $n = n_1 + n_2$. If H_0 is true, the numerator is

$$\frac{n_1 n_2}{n}(\bar{Y} - \bar{X}) = \sqrt{\frac{n_1 n_2}{n}}[\frac{\sigma_2}{\sqrt{n_2}}\sqrt{n_2}(\bar{Y} - \mu_2)/\sigma_2 - \frac{\sigma_1}{\sqrt{n_1}}\sqrt{n_1}(\bar{X} - \mu_1)/\sigma_1]$$

$$= \sigma_2\sqrt{n_1/n}Z_{2n} - \sigma_1\sqrt{n_2/n}Z_{1n},$$

- By the CLT,

$$\sigma_1 Z_{1n} \to^{\mathcal{L}} N(0, \sigma_1^2)$$

and

$$\sigma_2 Z_{2n} \to^{\mathcal{L}} N(0, \sigma_2^2)$$

and since, $n_1/n \to \lambda$ and $n_2/n = 1 - n_1/n \to 1 - \lambda$, then

$$\sigma_2\sqrt{n_1/n}Z_{2n} - \sigma_1\sqrt{n_2/n}Z_{1n} \to^{\mathcal{L}} N(0, \lambda\sigma_2^2 + (1-\lambda)\sigma_1^2)$$

since the X's and Y's are independent.

The denominator is

$$s = \sqrt{\frac{1}{n-2}[\sum_{i=1}^{n_1}(X_i - \bar{X})^2 + \sum_{i=1}^{n_2}(Y_i - \bar{Y})^2} = \sqrt{[\frac{n_1 - 1}{n-2}s_1^2 + \frac{n_1 - 1}{n}s_1^2]}$$

- By the consistency of plug-in estimates, $s_1^2 \to^{\mathcal{P}} \sigma_1^2$ and $s_2^2 \to^{\mathcal{P}} \sigma_2^2$.

- By results for functions (multi-dimensional case, B.7.1)

$$\sqrt{\frac{n_1 - 1}{n-2}s_1^2 + \frac{n_2 - 1}{n-2}s_1^2} \to^{\mathcal{P}} \sqrt{\lambda\sigma_1^2 + (1-\lambda)\sigma_2^2}$$

- Then, by Slutsky's Thereom (A.14.9) & results for functions

$$S_n \to^{\mathcal{L}} N(0, \frac{\lambda\sigma_2^2 + (1-\lambda)\sigma_1^2}{\lambda\sigma_1^2 + (1-\lambda)\sigma_2^2})$$