Do these problems in the spaces provided. Show your work.

1. (20 points) Let \(X\) and \(Y\) be independent and identically distributed (IID).
   
   (a) Show that if \(X\) and \(Y\) are continuous random variables, then \(P(X < Y) = 1/2\).

Let \(p\) be the common PDF for \(X\) and \(Y\). By independence,

\[
P(X < Y) = \sum_{x=-\infty}^{\infty} \left( \sum_{y=x}^{\infty} p(y) \, dy \right) \, p(x) \, dx
\]

\[
= \sum_{x=-\infty}^{\infty} (1 - F(x)) \, p(x) \, dx
\]

where \(F(x) = P(X < x)\) is the common CDF of \(X\) and \(Y\).

Since \(F' = p\), we have

\[
P(X < Y) = \sum_{-\infty}^{\infty} p(x) \, dx - \sum_{-\infty}^{\infty} F(x) \, p(x) \, dx
\]

\[
= 1 - \frac{1}{2} \left[ F(x) \right]_{-\infty}^{\infty}
\]

\[
= 1 - \frac{1}{2} \left( 1 - 0 \right)
\]

\[
= \frac{1}{2}
\]
(b) Give an example of two IID random variables $X$ and $Y$ such that $P(X < Y) \neq 1/2$.

Let $X$ and $Y$ be Bernoulli ($p$) and independent.

Then

$$P(X < Y) = P(X = 0, Y = 1)$$

$$= (1 - p) p$$

$$\neq \frac{1}{2}$$
2. (20 points) Recall that the beta function $B$ satisfies

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)},$$

where $\Gamma$ is the gamma function. Show that if $m$ and $n$ are positive integers, then

$$B(m, n) = \frac{(m - 1)!(n - 1)!}{(m + n - 1)!}$$

Since $\Gamma(n) = (n-1)!$ for all positive integers $n$, we have

$$B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} = \frac{(m-1)! (n-1)!}{(m+n-1)!}.$$
3. (20 points) Let $X$ be exponential with parameter $\lambda > 0$, and let $T$ be a continuous, nonnegative random variable, independent from $X$.

(a) Prove that, for any (non-random) $s > 0$,

$$P(X > s + T | X > T) = P(X > s)$$

We have

$$P(X > s + T | X > T) = \frac{P(X > s + T)}{P(X > T)}$$

Let $p_T$ be the PDF of $T$. Then

$$P(X > s + T) = \int_0^\infty P(X > s + t | T = t) p_T(t) \, dt$$

(by independence)

$$= \int_0^{\infty} P(X > s + t) p_T(t) \, dt$$

(law of total prob.)

$$= \int_0^{\infty} e^{-\lambda(s+t)} p_T(t) \, dt$$

$$= e^{-\lambda s} \int_0^{\infty} e^{-\lambda t} p_T(t) \, dt$$

$$= e^{-\lambda s} E(e^{-\lambda T})$$

So

$$P(X > s + T | X > T) = \frac{e^{-\lambda s} E(e^{-\lambda T})}{e^{\lambda s} E(e^{-\lambda T})} = e^{-\lambda s} = P(X > s)$$
(b) Use the previous result to prove that, if in addition $S$ is a continuous nonnegative random variable such that \{X, T, S\} is independent, then

$$P(X > S + T \mid X > T) = P(X > S)$$

(This is a “randomized” memoryless property of the exponential distribution.)

Let $f_S$ be the PDF of $S$. By total probability

$$P(X > S + T \mid X > T)$$

$$= \sum_{s=0}^{\infty} P(X > S + T \mid X > T, S=s) P_S(S=s \mid X > T) \, ds$$

$$= \sum_{s=0}^{\infty} P(X > s + T \mid X > T) P_S(s) \, ds$$  \hspace{1cm} \text{(by independence)}

$$= \sum_{s=0}^{\infty} e^{-\lambda s} P_S(s) \, ds$$ \hspace{1cm} \text{(vs. problem)}

$$= E(e^{-\lambda S})$$

But

$$P(X > S) = \sum_{s=0}^{\infty} P(X > S \mid S=s) P_S(s) \, ds$$

$$= \sum_{s=0}^{\infty} P(X > s) P_S(s) \, ds$$ \hspace{1cm} \text{(independence)}

$$= \sum_{s=0}^{\infty} e^{-\lambda s} P_S(s) \, ds = E(e^{-\lambda S})$$

So

$$P(X > S + T \mid X > T) = P(X > S)$$
4. (20 points) Recall the definition of the t-distribution: let \( U \) be \( \chi^2 \) with \( r \) degrees of freedom, let \( Z \) be standard normal, and let \( U \) and \( Z \) be independent. Then the random variable

\[
T = \frac{Z}{\sqrt{U/r}}
\]

has the t-distribution with \( r \) degrees of freedom. It can be shown that \( E[T] = 0 \) for \( r \geq 2 \). Therefore, the variance of \( T \) is

\[
\text{Var}(T) = E[T^2] = E\left[ \frac{Z^2}{U/r} \right]
\]

Find \( \text{Var}(T) \), justifying all computations.

By independence

\[
E\left( \frac{Z^2}{U/r} \right) = E(Z^2)E\left( \frac{1}{U} \right)
\]

The PDF of \( U \) is \( p(u) = \frac{1}{2^{r/2}\Gamma(r/2)} u^{r/2-1} e^{-u/2} \quad (u > 0) \)

So

\[
E\left( \frac{1}{U} \right) = \frac{1}{2^{r/2}\Gamma(r/2)} \int_{0}^{\infty} \frac{1}{u} \ u^{r/2-1} \ e^{-u/2} \ du
\]

\[
= \frac{1}{2^{r/2}\Gamma(r/2)} \int_{0}^{\infty} \ u^{r/2-2} \ e^{-u/2} \ du
\]

\[
= \frac{1}{2^{r/2}\Gamma(r/2)} \ 2^{r/2-1} \ \Gamma\left( \frac{r}{2} - 1 \right)
\]

\[
= \frac{1}{2} \ \frac{\Gamma\left( \frac{r}{2} - 1 \right)}{(r-1)\Gamma\left( \frac{r}{2} - 1 \right)} = \frac{1}{2} \ \frac{1}{\frac{r}{2} - 1} = \frac{\frac{r}{2}}{r-2}
\]

So \( \text{Var}(T) = \frac{r}{2} \ \frac{1}{\frac{r}{2} - 1} = \frac{r}{r-2} \)
5. (20 points) (Bayesian Poisson Estimation.) Suppose that $X$ is Poisson with mean $\lambda$, where $\lambda$ is unknown. To model our uncertainty in $\lambda$, we consider it a random variable $\Lambda$ whose prior distribution is gamma: $\Lambda \sim \Gamma(\alpha, \theta)$.

(a) The prior predictive distribution of $X$ is defined to be the distribution of a Poisson random variable with parameter $\Lambda$, where $\Lambda \sim \Gamma(\alpha, \theta)$. Find the prior predictive distribution of $X$.

\[
P(X = k) = \sum_{\lambda=0}^{\infty} P(X = k | \Lambda = \lambda) P_{\theta}(\Lambda = \lambda) \\lambda^k \frac{1}{k!} e^{-\lambda} \frac{\lambda^\alpha - 1}{\Gamma(\alpha)} \, d\lambda
\]
\[
= \frac{1}{k! \theta^\alpha \Gamma(\alpha)} \sum_{\lambda=0}^{\infty} \lambda^{k + \alpha - 1} \left(1 + \frac{\lambda}{\theta}\right) \, d\lambda
\]
\[
= \frac{1}{k! \theta^\alpha \Gamma(\alpha)} \left(1 + \frac{1}{\theta}\right)^{k + \alpha} \Gamma(k + \alpha)
\]
\[
= \frac{\Gamma(k + \alpha)}{k! \Gamma(\alpha)} \left(\frac{\theta}{\theta + 1}\right)^{k + \alpha} f_k(k = 0, 1, 2, \ldots)
\]

(b) Suppose that we observe a single value of $X$, say $X_1 = k_1$. Find the posterior distribution $P_\theta(\Lambda = \lambda | X_1 = k_1)$.

\[
P_\theta(\Lambda = \lambda | X_1 = k_1) = \frac{P(X_i = k_1 | \Lambda = \lambda) P_{\theta}(\Lambda = \lambda)}{\sum_{\lambda=0}^{\infty} P(X_i = k_1 | \Lambda = \lambda) P_{\theta}(\Lambda = \lambda) \, d\lambda}
\]
\[
= \frac{\theta^{k_1} \frac{\lambda^\alpha - 1}{\Gamma(\alpha)} \frac{1}{k!} e^{-\lambda} \frac{\lambda^\alpha - 1}{\Gamma(\alpha)} \, d\lambda}{\sum_{\lambda=0}^{\infty} e^{-\lambda} \frac{\lambda^\alpha - 1}{\Gamma(\alpha)} \frac{1}{k!} e^{-\lambda} \frac{\lambda^\alpha - 1}{\Gamma(\alpha)} \, d\lambda}
\]
\[
= e^{-\lambda} \left(1 + \frac{1}{\theta}\right)^{k_1 + \alpha - 1}
\]

where $C$ is independent of $\lambda$. So the posterior is $\Gamma(\alpha_{\text{post}} = \alpha + k_1, \theta_{\text{post}} = \frac{1}{1 + \frac{1}{\theta}})$.
(c) Use the result of the previous part to find the posterior distribution of $\Lambda$ when a sample mean of $\bar{X} = \bar{x}$ is observed from a sample of size $n$.

Consider how the posterior parameters change with each new observation $X_i = k_i$

$\alpha \mapsto \alpha + k_1 \mapsto \alpha + k_1 + k_2 \mapsto \ldots \mapsto \alpha + k_1 + k_2 + \ldots + k_n = \alpha + n\bar{x}$

$\theta \mapsto \frac{1}{1 + \frac{1}{\theta}} \mapsto \frac{1}{2 + \frac{1}{\theta}} \mapsto \ldots \mapsto \frac{1}{n + \frac{1}{\theta}}$

So the posterior after observing $\bar{X} = \bar{x}$ from a sample of size $n$ is

$$\mathcal{N}(\alpha_{\text{post}} = \alpha + n\bar{x}, \theta_{\text{post}} = \frac{1}{n + \frac{1}{\theta}})$$

(d) The posterior predictive distribution is defined to be the distribution of a Poisson random variable with parameter $\Lambda$, where $\Lambda$ comes from the posterior distribution. That is, the posterior predictive is the mass function $P(Y = k | \bar{X} = \bar{x})$. Find the posterior predictive distribution.

To ease notation, let $\alpha_p = \alpha_{\text{post}}, \theta_p = \theta_{\text{post}}$. Then

$$P(Y = k | \bar{X} = \bar{x}) = \sum_{\Lambda = 0}^{\infty} P(Y = k | \Lambda = \Lambda, \bar{X} = \bar{x}) P(\Lambda = \Lambda | \bar{X} = \bar{x}) \, d\Lambda$$

$$= \sum_{\Lambda = 0}^{\infty} e^{-\Lambda} \frac{\Lambda^k}{k!} \frac{1}{\theta_p} \frac{\Gamma(\alpha_p)}{\Gamma(\alpha_p + k + 1)} e^{-\Lambda} \theta_p \, d\Lambda$$

$$= \frac{1}{k! \theta_p \Gamma(\alpha_p)} \sum_{\Lambda = 0}^{\infty} \Lambda^k \theta_p \, d\Lambda$$

$$= \frac{1}{k! \theta_p \Gamma(\alpha_p)} \left( \frac{1}{1 + \frac{1}{\theta_p}} \right)^{k+\alpha_p} \Gamma(k+\alpha_p), \quad k: 0, 1, 2, \ldots$$
6. (20 points) A coin has unknown head probability $p$. You flip the coin $n$ times, and observe $X = k$ heads. Assuming a uniform prior for $p$, we have seen that the posterior distribution of $p$ is beta, with $\alpha = k + 1$, $\beta = n - k + 1$. Now you prepare to begin flipping the coin again. Consider the random variable

$Y =$ the number of additional flips required until the first head appears

(a) Find the conditional distribution of $Y$ given $p$; that is, $P(Y = j|p = \theta)$, for $j = 1, 2, 3, \ldots$

$$P(Y = j | p = \theta) = (1-\theta)^{j-1} \theta$$

for $j = 1, 2, 3, \ldots$

(I.e. given $p = \theta$, $Y$ is Geo ($p=\theta$))
(b) Find the posterior predictive distribution of $Y$ given $X = k$; that is, $P(Y = j|X = k)$, for $j = 1, 2, 3, \ldots$

\[
P(Y = j | X = k) = \sum_{\theta = 0}^{1} P(Y = j | p = \theta, X = k) P(p = \theta | X = k) \, d\theta
\]

\[
= \sum_{\theta = 0}^{1} (1-\theta)^{n-k} \theta \frac{1}{B(k+1, n-k+1)} \theta^k (1-\theta)^{n-k} \, d\theta
\]

\[
= \frac{1}{B(k+1, n-k+1)} \sum_{\theta = 0}^{1} \theta^{k+1} (1-\theta)^{n-k+j-1} \, d\theta
\]

\[
= \frac{B(k+2, n-k+j)}{B(k+1, n-k+1)} \quad \text{for } j = 1, 2, 3, \ldots
\]