# Take Home Exam (Spring Term 2016)

1. (a) Let us first recall the two required results (equation 1.39 and example A.79):

$$\Gamma(a) = 2 \int_0^\infty u^{2a-1} e^{-u^2} \, \mathrm{d}u \qquad (1.39)$$

$$\sqrt{\pi} = \int_{-\infty}^\infty e^{-x^2} \, \mathrm{d}x \qquad (A.79)$$

Additionally, it is helpful to establish that the function $f(x) = e^{-x^2}$ is even. Recall that a function is even if the following equation holds for all values contained in the domain:

$$f(-x) = f(x)$$

Using the function $f(x) = e^{-x^2}$:

$$\underbrace{e^{-(-x)^2}}_{=f(-x)} = \underbrace{e^{-x^2}}_{=f(x)}$$

And hence we can conclude that the function in question is indeed even. We will use this result later in our proof.

**Proposition:** $\qquad \Gamma(1/2) = \sqrt{\pi}$

*Proof.* In order to prove that $\Gamma(1/2)$ equals $\sqrt{\pi}$, we can simply plug $\frac{1}{2}$ into our equation from 1.39 and start simplifying the resulting expression.

$$\Gamma(a) = 2 \int_0^\infty u^{2a-1} e^{-u^2} \, \mathrm{d}u$$

Substituting in $\frac{1}{2}$ for $a$:

$$\Gamma(1/2) = 2 \int_0^\infty u^{2\frac{1}{2}-1} e^{-u^2} \, \mathrm{d}u$$

$$= 2 \int_0^\infty e^{-u^2} \, \mathrm{d}u$$

Due to the obviously symmetrical nature of the integrand $e^{-u^2}$ (it is an even function as discussed at the beginning of this proof), we can rearrange the expression as follows:

$$= \int_{-\infty}^\infty e^{-u^2} \, \mathrm{d}u$$

It is now trivial to see that we could just use $x$ instead of $u$ as our variable of choice.

$$= \int_{-\infty}^\infty e^{-x^2} \, \mathrm{d}x$$

This is the exact same expression used in example A.79 and hence we can apply it here, which yields our end result.

$$\Gamma(1/2) = \sqrt{\pi}$$

∎

(b) The second approach makes use of the following two equations:

$$B(a,b) = 2 \int_0^{\pi/2} (\sin\theta)^{2a-1}(\cos\theta)^{2b-1}\,\mathrm{d}\theta \tag{1}$$

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{2}$$

Again, it is helpful to first show a different result, which in turn will facilitate our actual proof.

**Proposition:** $\Gamma(1) = 1$

*Proof.*

$$\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}\,\mathrm{d}x$$

Setting $a = 1$:

$$\begin{aligned}
\Gamma(1) &= \int_0^\infty x^{1-1}e^{-x}\,\mathrm{d}x = \int_0^\infty x^0 e^{-x}\,\mathrm{d}x \\
&= \int_0^\infty e^{-x}\,\mathrm{d}x \\
&= \left[-e^{-x}\right]_0^\infty \\
&= \lim_{x\to\infty} -e^{-x} + e^0 \\
&= 0 + 1 \\
&= 1
\end{aligned}$$

∎

We're now able to do the actual proof.

**Proposition:** $\Gamma(1/2) = \sqrt{\pi}$

*Proof.* Let us start by combining our two initial equations (1) and (2):

$$\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = 2\int_0^{\pi/2}(\sin\theta)^{2a-1}(\cos\theta)^{2b-1}\,\mathrm{d}\theta$$

Setting $a = b = \frac{1}{2}$ yields:

$$\frac{\Gamma(1/2)\Gamma(1/2)}{\Gamma(1)} = 2\int_0^{\pi/2}(\sin\theta)^{2\frac{1}{2}-1}(\cos\theta)^{2\frac{1}{2}-1}\,\mathrm{d}\theta$$

$$\Longleftrightarrow \frac{(\Gamma(1/2))^2}{\Gamma(1)} = 2\int_0^{\pi/2} (\sin\theta)^{1-1}(\cos\theta)^{1-1}\,\mathrm{d}\theta$$

$$= 2\int_0^{\pi/2}\mathrm{d}\theta$$

As shown earlier, $\Gamma(1)$ simplifies to 1, and hence our equation is reduced to:

$$(\Gamma(1/2))^2 = 2\int_0^{\pi/2}\mathrm{d}\theta$$

$$= 2\cdot\theta|_0^{\pi/2}$$

$$= 2(\frac{\pi}{2} - 0)$$

$$= \pi$$

$$\Longleftrightarrow \Gamma(1/2) = \pm\sqrt{\pi}$$

Given the definition of the gamma function, only the positive root is possible and hence we've shown:

$$\Gamma(1/2) = \sqrt{\pi}$$

∎

2. (a) In order to demonstrate an approach that allows us to evaluate the definite integral $I = \int_0^2 y^4(2-y)^3\,\mathrm{d}y$ without using a calculator, it is helpful to first show two different results: the gamma difference equation and the relationship between the gamma function and the factorial function. Let us start with the gamma difference equation, which states the following:

**Proposition:** $\qquad \Gamma(a) = (a-1)\Gamma(a-1) \quad \forall a \in \mathbb{R}_{>1}$

*Proof.*

$$\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}\,\mathrm{d}x$$

Let us now apply integration by parts with $u = x^{a-1}$ and $\mathrm{d}v = e^{-x}\,\mathrm{d}x$. This yields $\mathrm{d}u = (a-1)x^{a-2}\,\mathrm{d}x$ and $v = -e^{-x}$. Hence:

$$\Gamma(a) = u\cdot v|_{x=0}^{\infty} - \int_0^\infty v\,\mathrm{d}u$$

$$= \underbrace{\left[-e^{-x}x^{a-1}\right]_{x=0}^{\infty}}_{=0} + \int_0^\infty e^{-x}(a-1)x^{(a-1)-1}\,\mathrm{d}x$$

$$= (a-1)\int_0^\infty e^{-x}x^{(a-1)-1}\,\mathrm{d}x$$

$$= (a-1)\Gamma(a-1)$$

∎

The second useful proposition we will need for our upcoming proof describes the gamma function as an extension of the factorial function.

**Proposition:** $\Gamma(n) = (n-1)! \quad \forall n \in \mathbb{N}$

*Proof.* This problem lends itself well to an induction-based proof.

**Base case:** The base case in this situation is $n = 1$, which means we need to show that $\Gamma(1) = 0!$. Since $0!$ is equal to one and we have already shown that $\Gamma(1) = 1$ during the previous question, the base case holds true.

**Induction step:** Assuming that our proposition holds for $n = k$, we now need to show that it also holds for $n = k + 1$. The easiest way to do so consists of applying the gamma difference equation, which we proved above. According to this result, we can rewrite $\Gamma(k+1)$ as follows:

$$\Gamma(k+1) = k\Gamma(k)$$

Since we're assuming the relationship holds for $n = k$:

$$= k(k-1)!$$
$$= k!$$

The proposition thus follows by mathematical induction.

■

We're now ready to show the actual result in question without relying on a calculator.

**Proposition:** $I = \int_0^2 y^4(2-y)^3 \, \mathrm{d}y = \frac{32}{35}$

*Proof.* Our goal is to transform the expression into a beta function. This can be accomplished using the substitution $u = \frac{y}{2}$, which implies $y = 2u$, $y^4 = 2^4 u^4$, and $\mathrm{d}y = 2\,\mathrm{d}u$:

$$I = \int_0^2 y^4(2-y)^3 \, \mathrm{d}y$$
$$= \int_0^1 2^4 u^4(2-2u)^3 2 \, \mathrm{d}u$$
$$= \int_0^1 2^5 u^4(2(1-u))^3 \, \mathrm{d}u$$
$$= \int_0^1 2^5 u^4 2^3(1-u)^3 \, \mathrm{d}u$$
$$= 2^8 \int_0^1 u^4(1-u)^3 \, \mathrm{d}u$$
$$= 2^8 \int_0^1 u^{5-1}(1-u)^{4-1} \, \mathrm{d}u$$
$$= 2^8 B(5,4)$$

Using the relationship between the gamma and beta functions aids us in our efforts to further simplify the expression.

$$I = 2^8 \frac{\Gamma(5)\Gamma(4)}{\Gamma(5+4)} = 2^8 \frac{\Gamma(5)\Gamma(4)}{\Gamma(9)}$$

We can now make use of the relationship between the gamma function and the factorial function we proved earlier.

$$\begin{aligned}
I &= 2^8 \frac{4! \cdot 3!}{8!} \\
&= 2^8 \frac{4 \cdot 3 \cdot 2 \cdot 3 \cdot 2}{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2} \\
&= 2^8 \frac{6}{8 \cdot 7 \cdot 6 \cdot 5} \\
&= 2^5 2^3 \frac{6}{8 \cdot 7 \cdot 6 \cdot 5} \\
&= \frac{2^5}{7 \cdot 5} \\
&= \frac{32}{35}
\end{aligned}$$

$\blacksquare$

(b) **Proposition:** $\qquad B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} \, \mathrm{d}t = \int_0^\infty s^{a-1} \left(\frac{1}{s+1}\right)^{a+b} \, \mathrm{d}s$

*Proof.* As indicated by the question statement, we need to find another integral substitution to show the proposed alternate formulation of the beta function. Let us start by substituting $s = \frac{t}{1-t}$. We also need to be able to express $t$ in terms of s, so let's derive an expression for that as well (based on our substitution):

$$\begin{aligned}
& s = \frac{t}{1-t} \\
\Longleftrightarrow \quad & \frac{1-t}{t} = \frac{1}{s} \\
\Longleftrightarrow \quad & \frac{1}{t} - 1 = \frac{1}{s} \\
\Longleftrightarrow \quad & \frac{1}{t} = \frac{s+1}{s} \\
\Longleftrightarrow \quad & t = \frac{s}{s+1}
\end{aligned}$$

Additionally, using the quotient rule for derivatives, we can derive:

$$\begin{aligned}
\frac{\mathrm{d}t}{\mathrm{d}s} &= \frac{(s+1) - s}{(s+1)^2} = \frac{1}{(s+1)^2} \\
\Longleftrightarrow \quad \mathrm{d}t &= \frac{1}{(s+1)^2} \, \mathrm{d}s
\end{aligned}$$

Given our substitution, the bounds of integration obviously change as well, which is what we were looking for. Let us now start with the original integral (the classic beta function) and apply our substitution:

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} \, dt$$

$$= \int_0^\infty \left(\frac{s}{s+1}\right)^{a-1} \left(1 - \frac{s}{s+1}\right)^{b-1} \frac{1}{(s+1)^2} \, ds$$

$$= \int_0^\infty \left(\frac{s}{s+1}\right)^{a-1} \left(\frac{s+1}{s+1} - \frac{s}{s+1}\right)^{b-1} \frac{1}{(s+1)^2} \, ds$$

$$= \int_0^\infty \left(\frac{s}{s+1}\right)^{a-1} \left(\frac{1}{s+1}\right)^{b-1} \left(\frac{1}{s+1}\right)^2 \, ds$$

$$= \int_0^\infty \left(\frac{s}{s+1}\right)^{a-1} \left(\frac{1}{s+1}\right)^{b+1} \, ds$$

$$= \int_0^\infty s^{a-1} \left(\frac{1}{s+1}\right)^{a-1} \left(\frac{1}{s+1}\right)^{b+1} \, ds$$

$$= \int_0^\infty s^{a-1} \left(\frac{1}{s+1}\right)^{a-1+b+1} \, ds$$

$$= \int_0^\infty s^{a-1} \left(\frac{1}{s+1}\right)^{a+b} \, ds$$

∎

(c) **Proposition:** $\int_0^\infty e^{-mt} t^{z-1} \, dt = \frac{\Gamma(z)}{m^z}$

*Proof.*

Substituting $u = mt$ leads to $t = \frac{u}{m}$, $dt = \frac{1}{m} \, du$, and thus:

$$\int_0^\infty e^{-mt} t^{z-1} \, dt = \int_0^\infty e^{-u} \left(\frac{u}{m}\right)^{z-1} \frac{1}{m} \, du$$

$$= \int_0^\infty e^{-u} u^{z-1} \left(\frac{1}{m}\right)^{z-1} \frac{1}{m} \, du$$

$$= \frac{1}{m^z} \int_0^\infty e^{-u} u^{z-1} \, du$$

$$= \frac{\Gamma(z)}{m^z}$$

∎

Let us now prove the relationship between the gamma and beta function as requested by the exam question.

**Proposition:** $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$

*Proof.*

$$\int_0^\infty e^{-mt}t^{z-1} \, \mathrm{d}t = \frac{\Gamma(z)}{m^z}$$

Applying the proposed substitution ($m = s + 1$ and $z = x + y$):

$$\int_0^\infty e^{-(s+1)t}t^{x+y-1} \, \mathrm{d}t = \frac{\Gamma(x+y)}{(s+1)^{x+y}}$$

$$\Gamma(x+y) = (s+1)^{x+y} \int_0^\infty e^{-(s+1)t}t^{x+y-1} \, \mathrm{d}t$$

Since we're interested in the relationship between the beta and the gamma function, let us incorporate the beta function:

$$B(x,y)\Gamma(x+y) = B(x,y)(s+1)^{x+y} \int_0^\infty e^{-(s+1)t}t^{x+y-1} \, \mathrm{d}t$$

We are now going to use (on the right hand side of our equation) the alternate formulation of the beta function we encountered earlier. Then, based on Tonelli's theorem, which applies in this case due to the positive nature of the integrand(s), we obtain:

$$B(x,y) \cdot \Gamma(x+y) = \int_0^\infty \int_0^\infty s^{x-1}\left(\frac{1}{s+1}\right)^{x+y} \cdot (s+1)^{x+y}e^{-(s+1)t}t^{x+y-1} \, \mathrm{d}t \, \mathrm{d}s$$

$$= \int_0^\infty \int_0^\infty s^{x-1}e^{-(s+1)t}t^{x+y-1} \, \mathrm{d}t \, \mathrm{d}s$$

$$= \int_0^\infty \int_0^\infty s^{x-1}e^{-(s+1)t}t^{x-1}t^y \, \mathrm{d}t \, \mathrm{d}s$$

$$= \int_0^\infty \int_0^\infty (st)^{x-1}e^{-(s+1)t}t^y \, \mathrm{d}t \, \mathrm{d}s$$

Substituting $u = st$ leads to $s = \frac{u}{t}$, $\mathrm{d}s = \frac{1}{t} \, \mathrm{d}u$, and thus:

$$B(x,y) \cdot \Gamma(x+y) = \int_0^\infty \int_0^\infty u^{x-1}e^{-(\frac{u}{t}+1)t}t^y\frac{1}{t} \, \mathrm{d}t \, \mathrm{d}u$$

$$= \int_0^\infty \int_0^\infty u^{x-1}e^{-u}e^{-t}t^yt^{-1} \, \mathrm{d}t \, \mathrm{d}u$$

$$= \int_0^\infty \int_0^\infty u^{x-1}e^{-u}e^{-t}t^{y-1} \, \mathrm{d}t \, \mathrm{d}u$$

Applying Tonelli's theorem once more yields:

$$B(x,y) \cdot \Gamma(x+y) = \int_0^\infty u^{x-1}e^{-u} \, \mathrm{d}u \cdot \int_0^\infty e^{-t}t^{y-1} \, \mathrm{d}t$$

$$= \Gamma(x)\Gamma(y)$$

$$\iff B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

∎

3. (a) Binomial theorem:

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Vandermonde's theorem:

$$\binom{N + M}{k} = \sum_{i=0}^{k} \binom{N}{i} \binom{M}{k - i}, \quad 0 < k \leq \min(N, M)$$

(b) Let us assume an urn with $N$ white balls and $M$ black balls. We then sample $n$ balls (without replacement) from said urn. If we then denote the number of white balls drawn as $X$, we've defined a random variable with a hypergeometric distribution. $X$ then has the following pmf:

$$f_{\text{HGeo}}(x; N, M, n) = \frac{\binom{N}{x} \binom{M}{n-x}}{\binom{N+M}{n}} \mathbb{I}_{\{\max(0, n-M), \ldots, \min(n, N)\}}(x)$$

The above pmf assigns probabilities to various potential outcomes. More precisely, $f_{\text{HGeo}}(x; N, M, n)$ yields the probability of drawing exactly $x$ white balls and $n - x$ black balls. In line with common intuition, we can interpret the numerator of the pmf as the number of ways we could obtain $x$ white balls and $n-x$ black balls. As there are $N$ white balls, there exist $\binom{N}{x}$ combinations with exactly $x$ white balls. Since we are drawing $n$ balls in total, $n - x$ are still to be withdrawn from the urn. Recall that there are $M$ black balls and thus there are $\binom{M}{n-x}$ combinations with exactly $n - x$ black balls. This lets us conclude that there must be $\binom{N}{x} \binom{M}{n-x}$ combinations that result in $x$ white balls and $n - x$ black balls. The denominator on the other hand contains an expression for the total amount of outcomes and should be fairly self-explanatory. Given the $N$ white balls and the $M$ back balls, there are $N + M$ balls contained within the urn. Since we're picking $n$ balls, the total amount of outcomes is just $\binom{N+M}{n}$. So the fraction does indeed seem to represent the appropriate probability. Now all that's left to explain is the indicator function, which might look somewhat daunting at first, but is actually quite easy to understand. First of all, no matter the exact values of $N, M$, and $n$, $x$ can't be negative and it can't exceed $N$ either (it's not possible to draw more white balls than there actually are since we're sampling without replacement). If $n - M > 0$ - that is to say we draw more balls than there are black balls - the difference of $n$ and $M$ constitutes a lower bound as well. Even if we picked all $M$ of the black balls, we'd still have to sample $n - M$ more times and would thus pick $n - M$ white balls since there are no more black ones left. We've now explained the second part of the maximum function found in the indicator function. Let us now shift our focus to the upper bound, which involves a minumum function. As we've already mentioned, $x$ obviously cannot be larger than $N$ since there are only $N$ white balls. Furthermore, we cannot draw more white balls than actual draws and so the upper bound should be clear as well.

Let us now spend some time on a few asymptotic arguments. It should be obvious that the hypergeometric distribution resembles certain aspects of the binomial distribution since they only differ in one key point. While the binomial distribution describes sampling **with** replacement, the hypergeometric distribution considers the case of sampling **without** replacement. Intuitively, it should be easy to digest that the difference between

the binomial and hypergeometric distributions approaches zero as $N$ and $M$ approach infinity while we keep the ratio of the two and our sample size $n$ constant. Let us now provide a formal proof for this claim.

**Proposition:** Let $X$ be a discrete random variable with a hypergeometric distribution such that $X \sim \text{HGeo}(N, M, n)$. Keeping sample size $(n)$ and the proportion of white balls (denoted as $p$) constant:

$$\lim_{\substack{N \to \infty \\ M \to \infty}} f_X(x; N, M, n) = \underbrace{\binom{n}{x} p^x (1 - p)^{n-x} \mathbb{I}_{\{0,\dots,n\}}(x)}_{\text{pmf of Bin}(n,p)}$$

*Proof.* Let us start by rearranging the pmf of $X$ we stated at the beginning of this sub-question. First of all - as a means of improving readability - we'll denote the subscript of the indicator function of the hypergeometric (the set) as $S$ for the majority of this proof and only tackle it right at the end as it's extremely easy to handle.

$$f_X(x; N, M, n) = \frac{\binom{N}{x}\binom{M}{n-x}}{\binom{N+M}{n}} \mathbb{I}_S(x)$$

$$= \frac{N!}{x!(N-x)!} \frac{M!}{(n-x)!(M-(n-x))!} \frac{n!(N+M-n)!}{(N+M)!} \mathbb{I}_S(x)$$

$$= \binom{n}{x} \frac{N!}{(N-x)!} \frac{M!}{(M-(n-x))!} \frac{(N+M-n)!}{(N+M)!} \mathbb{I}_S(x)$$

$$= \binom{n}{x} N(N-1)\dots(N-x+1) \frac{M!}{(M-(n-x))!} \frac{(N+M-n)!}{(N+M)!} \mathbb{I}_S(x)$$

$$= \binom{n}{x} N(N-1)\dots(N-x+1) M(M-1)\dots(M-(n-x)+1)$$

$$\frac{(N+M-n)!}{(N+M)!} \mathbb{I}_S(x)$$

$$= \binom{n}{x} N(N-1)\dots(N-x+1) M(M-1)\dots(M-(n-x)+1)$$

$$\frac{1}{(N+M)(N+M-1)\dots(N+M-n+1)} \mathbb{I}_S(x)$$

$$= \binom{n}{x} \frac{N(N-1)\dots(N-x+1) M(M-1)\dots(M-(n-x)+1)}{(N+M)(N+M-1)\dots(N+M-n+1)} \mathbb{I}_S(x)$$

$$= \binom{n}{x} \prod_{i=1}^{x} \frac{N-x+i}{M+N-x+i} \cdot \prod_{j=1}^{n-x} \frac{M-(n-x)+j}{M+N-n+j} \mathbb{I}_S(x)$$

Let us now study the asymptotic behavior of this pmf as $N$ and $M$ approach infinity, while the proportion of white balls $p = \frac{N}{N+M}$ remains constant.

$$\lim_{\substack{N \to \infty \\ M \to \infty}} f_X(x; N, M, n) = \lim_{\substack{N \to \infty \\ M \to \infty}} \binom{n}{x} \underbrace{\prod_{i=1}^{x} \frac{N-x+i}{M+N-x+i}}_{\equiv \mathcal{Q}_1} \cdot \underbrace{\prod_{j=1}^{n-x} \frac{M-(n-x)+j}{M+N-n+j}}_{\equiv \mathcal{Q}_2} \mathbb{I}_S(x)$$

Assuming the limits of all these factors exist, we can express the limit of the product as the product of limits:

$$\lim_{\substack{N\to\infty\\M\to\infty}} f_X(x;N,M,n) = \lim_{\substack{N\to\infty\\M\to\infty}} \binom{n}{x} \cdot \lim_{\substack{N\to\infty\\M\to\infty}} \mathcal{Q}_1 \cdot \lim_{\substack{N\to\infty\\M\to\infty}} \mathcal{Q}_2 \cdot \lim_{\substack{N\to\infty\\M\to\infty}} \mathbb{I}_S(x)$$

$$= \binom{n}{x} \cdot \lim_{\substack{N\to\infty\\M\to\infty}} \mathcal{Q}_1 \cdot \lim_{\substack{N\to\infty\\M\to\infty}} \mathcal{Q}_2 \cdot \lim_{\substack{N\to\infty\\M\to\infty}} \mathbb{I}_S(x) \tag{3}$$

It is now necessary to compute these three limits contained in equation (3).

$$\lim_{\substack{N\to\infty\\M\to\infty}} \mathcal{Q}_1 = \lim_{\substack{N\to\infty\\M\to\infty}} \prod_{i=1}^{x} \frac{N-x+i}{M+N-x+i}$$

$$= \prod_{i=1}^{x} \lim_{\substack{N\to\infty\\M\to\infty}} \frac{N-x+i}{M+N-x+i}$$

$$= \prod_{k=1}^{x} \lim_{\substack{N\to\infty\\M\to\infty}} \frac{\frac{N}{N+M} + \frac{i-x}{N+M}}{1 + \frac{i-x}{N+M}}$$

$$= \prod_{k=1}^{x} \lim_{\substack{N\to\infty\\M\to\infty}} \frac{p + \frac{i-x}{N+M}}{1 + \frac{i-x}{N+M}}$$

$$= \prod_{k=1}^{x} p$$

$$= p^x \tag{4}$$

$$\lim_{\substack{N\to\infty\\M\to\infty}} \mathcal{Q}_2 = \lim_{\substack{N\to\infty\\M\to\infty}} \prod_{j=1}^{n-x} \frac{M-(n-x)+j}{M+N-n+j}$$

$$= \lim_{\substack{N\to\infty\\M\to\infty}} \prod_{j=1}^{n-x} \frac{M-n+x+j}{M+N-n+j}$$

$$= \prod_{j=1}^{n-x} \lim_{\substack{N\to\infty\\M\to\infty}} \frac{M-n+x+j}{M+N-n+j}$$

$$= \prod_{j=1}^{n-x} \lim_{\substack{N\to\infty\\M\to\infty}} \frac{\frac{M}{N+M} + \frac{x+j-n}{N+M}}{1 + \frac{j-n}{N+M}}$$

$$= \prod_{j=1}^{n-x} \lim_{\substack{N\to\infty\\M\to\infty}} \frac{(1-p) + \frac{x+j-n}{N+M}}{1 + \frac{j-n}{N+M}}$$

$$= \prod_{j=1}^{n-x} (1-p)$$

$$= (1-p)^{n-x} \tag{5}$$

$$\lim_{\substack{N\to\infty\\M\to\infty}} \mathbb{I}_S(x) = \lim_{\substack{N\to\infty\\M\to\infty}} \mathbb{I}_{\{\max(0,n-M),\ldots,\min(n,N)\}}(x)$$

$$= \mathbb{I}_{\{0,\ldots,n\}}(x) \tag{6}$$

We can now insert our results from equations (4), (5), and (6) into equation (3):

$$\lim_{\substack{N\to\infty \\ M\to\infty}} f_X(x; N, M, n) = \binom{n}{x} \cdot \lim_{\substack{N\to\infty \\ M\to\infty}} \mathcal{Q}_1 \cdot \lim_{\substack{N\to\infty \\ M\to\infty}} \mathcal{Q}_2 \cdot \lim_{\substack{N\to\infty \\ M\to\infty}} \mathbb{I}_S(x)$$

$$= \underbrace{\binom{n}{x} p^x (1-p)^{n-x} \, \mathbb{I}_{\{0,\dots,n\}}(x)}_{\text{pmf of Bin}(n,p)}$$

We have thus shown that the limiting case (assuming certain conditions hold true) of the hypergeometric is a binomial with $p = \frac{N}{N+M}$. ∎

(c) We're first asked to derive the Poisson pmf as a limiting case of the binomial distribution.

**Proposition:** Let $Y \sim \text{Bin}(n, p)$ and $p = \lambda/n$. Then, as $n$ approaches infinity (and $p$ approaches zero):

$$\lim_{\substack{n\to\infty \\ p\to 0}} f_Y(y; n, p) = \underbrace{\frac{e^{-\lambda}\lambda^y}{y!} \, \mathbb{I}_{\{0,1,\dots\}}(y)}_{\text{pmf of Poi}(\lambda)}$$

*Proof.* Let us start by rearranging the pmf of $Y$.

$$f_Y(y; n, p) = \binom{n}{y} p^y (1-p)^{n-y} \, \mathbb{I}_{\{0,1,\dots,n\}}(y)$$

$$= \frac{n!}{(n-y)!\, y!} \left(\frac{\lambda}{n}\right)^y \left(1 - \frac{\lambda}{n}\right)^{n-y} \mathbb{I}_{\{0,1,\dots,n\}}(y)$$

$$= \frac{n!}{(n-y)!} \frac{1}{y!} \frac{\lambda^y}{n^y} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-y} \mathbb{I}_{\{0,1,\dots,n\}}(y)$$

$$= \frac{n(n-1)\dots(n-y+1)}{n^y} \frac{\lambda^y}{y!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-y} \mathbb{I}_{\{0,1,\dots,n\}}(y)$$

Let us now see what happens in the limit(s):

$$\lim_{\substack{n\to\infty \\ p\to 0}} f_Y(y; n, p) = \lim_{\substack{n\to\infty \\ p\to 0}} \frac{n(n-1)\dots(n-y+1)}{n^y} \frac{\lambda^y}{y!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-y} \mathbb{I}_{\{0,1,\dots,n\}}(y)$$

Assuming the limits of all these factors exist, we can express the limit of the product as the product of limits:

$$\lim_{\substack{n\to\infty \\ p\to 0}} f_Y(y; n, p) = \lim_{\substack{n\to\infty \\ p\to 0}} \frac{n(n-1)\dots(n-y+1)}{n^y} \cdot \lim_{\substack{n\to\infty \\ p\to 0}} \frac{\lambda^y}{y!} \cdot \lim_{\substack{n\to\infty \\ p\to 0}} \left(1 - \frac{\lambda}{n}\right)^n$$

$$\cdot \lim_{\substack{n\to\infty \\ p\to 0}} \left(1 - \frac{\lambda}{n}\right)^{-y} \cdot \lim_{\substack{n\to\infty \\ p\to 0}} \mathbb{I}_{\{0,1,\dots,n\}}(y)$$

Since the first limit expression's numerator multiplies $y$ expressions containing an $n$ and the denominator is just $n^y$, the entire first limit is obviously equal to 1.

$$\lim_{\substack{n\to\infty \\ p\to 0}} f_Y(y; n, p) = \underbrace{\lim_{\substack{n\to\infty \\ p\to 0}} \frac{\lambda^y}{y!}}_{=\frac{\lambda^y}{y!}} \cdot \underbrace{\lim_{\substack{n\to\infty \\ p\to 0}} \left(1 - \frac{\lambda}{n}\right)^n}_{=e^{-\lambda}} \cdot \underbrace{\lim_{\substack{n\to\infty \\ p\to 0}} \left(1 - \frac{\lambda}{n}\right)^{-y}}_{=1} \cdot \underbrace{\lim_{\substack{n\to\infty \\ p\to 0}} \mathbb{I}_{\{0,1,\dots,n\}}(y)}_{\mathbb{I}_{\{0,1,\dots\}}(y)}$$

$$= \frac{e^{-\lambda}\lambda^y}{y!} \, \mathbb{I}_{\{0,1,\dots\}}(y)$$

$\blacksquare$

Since we've shown (preceding sub-question) that the hypergeometric approaches (as $N \to \infty$, $M \to \infty$, and $p = const$) the binomial, and that the binomial approaches (as $n \to \infty$, and $p \to 0$) the Poisson, we can approximate the hypergeometric using a Poisson with $\lambda = np = \frac{nN}{N+M}$.

4. (a) **Proposition:** Let $Y$ be a continuous random variable with density $f_Y(y)$ symmetric about zero and cdf $F_Y$. Then:

$$\mathbb{E}\left[ F_Y\left( \frac{Y}{\sigma} \right) \right] = 0.5 \quad \forall \sigma > 0$$

*Proof.* Let $Z$ be another random variable, which is independent of $Y$ but follows the same distribution and thus has the same cdf ($F_Z(z) = F_Y(y)$).

$$F_Y(y) = \Pr(Y < y)$$

Since $Y$ and $Z$ follow the same distribution (by definition from above):

$$F_Y(y) = \Pr(Z < y)$$

Hence, we can express $F_Y\left(\frac{Y}{\sigma}\right)$ as follows:

$$F_Y\left( \frac{Y}{\sigma} \right) = \Pr\left( Z < \frac{Y}{\sigma} \right)$$
$$= \Pr\left( -Z > -\frac{Y}{\sigma} \right)$$
$$\overset{\text{sym.}}{=} \Pr\left( Z > \frac{Y}{\sigma} \right)$$

Note that this last step hinges on the fact that $Y$ (and thus $Z$) are symmetric about zero. Using basic properties from probability theory, we can rewrite our equation as follows.

$$F_Y\left( \frac{Y}{\sigma} \right) = 1 - \Pr\left( Z < \frac{Y}{\sigma} \right)$$
$$= 1 - F_Z\left( \frac{Y}{\sigma} \right)$$

Since $Y$ and $Z$ have the same cdf:

$$F_Y\left( \frac{Y}{\sigma} \right) = 1 - F_Y\left( \frac{Y}{\sigma} \right)$$
$$\iff 2 \cdot F_Y\left( \frac{Y}{\sigma} \right) = 1$$

$$\Longleftrightarrow F_Y\left(\frac{Y}{\sigma}\right) = 0.5$$

Since this is just a constant, taking its expectation has no effect.

$$\mathbb{E}\left[F_Y\left(\frac{Y}{\sigma}\right)\right] = \mathbb{E}\left[0.5\right]$$
$$= 0.5$$

We have thus shown the proposition. ∎

(b) Let us define $X = F_Y(Y/\sigma)$. We're now asked to derive an expression for the pdf of $X$. Note that following derivations in this question (and the ones yet to come) use the infimum definition (generalized inverse) of the inverse cdf (quantile function) as cdfs are - in general - weakly monotonic and right-continuous.

**Proposition:** $\qquad f_X(t; \sigma) = \sigma f_Y(\sigma F_Y^{-1}(t))\frac{\partial F_Y^{-1}(t)}{\partial t} \quad \forall \sigma > 0$

*Proof.* Let us first derive the cdf $F_X(x)$ of $X$.

$$F_X(x; \sigma) = \Pr(X \le x)$$
$$= \Pr\left(F_Y\left(\frac{Y}{\sigma}\right) \le x\right)$$
$$= \Pr\left(F_Y^{-1}\left(F_Y\left(\frac{Y}{\sigma}\right)\right) \le F_Y^{-1}(x)\right)$$
$$= \Pr\left(\frac{Y}{\sigma} \le F_Y^{-1}(x)\right)$$

We can multiply both sides of the inequality by $\sigma$ without changing the direction of the inequality as $\sigma$ is assumed to be positive.

$$F_X(x; \sigma) = \Pr\left(Y \le \sigma F_Y^{-1}(x)\right)$$
$$= F_Y(\sigma F_Y^{-1}(x))$$

Taking derivatives of both sides and applying the chain rule on the right hand side yields the desired result:

$$f_X(t; \sigma) = \sigma f_Y(\sigma F_Y^{-1}(t))\frac{\partial F_Y^{-1}(t)}{\partial t}, \quad \forall \sigma > 0$$

∎

(c) In order to create the requested comparison, I've written a small Python script that should be fairly self-explanatory as all important parts and lines are commented. Basically, the script simply simulates $10^6$ $Y$ values from a Student's $t$ distribution with $\nu = 3$ degrees of freedom. It then transforms these $10^6$ $Y$ values with the cdf of $Y$ to illustrate the classic probability integral transform (PIT) (not requested by the question statement but a nice illustration). Additionally, based on these $10^6$ realizations of $Y$, we compute $10^6$ realizations of $X$. We then apply kernel density estimation (KDE) in

order to implement the required comparison. Finally, the best-fitted beta distribution is estimated and it's pdf plotted. The results can be seen in the figure below. Notice that $X$ is not beta-distributed in general, as evidenced by the sharper peak of the KDE for $X$ (as opposed to the best-fitting beta distribution).
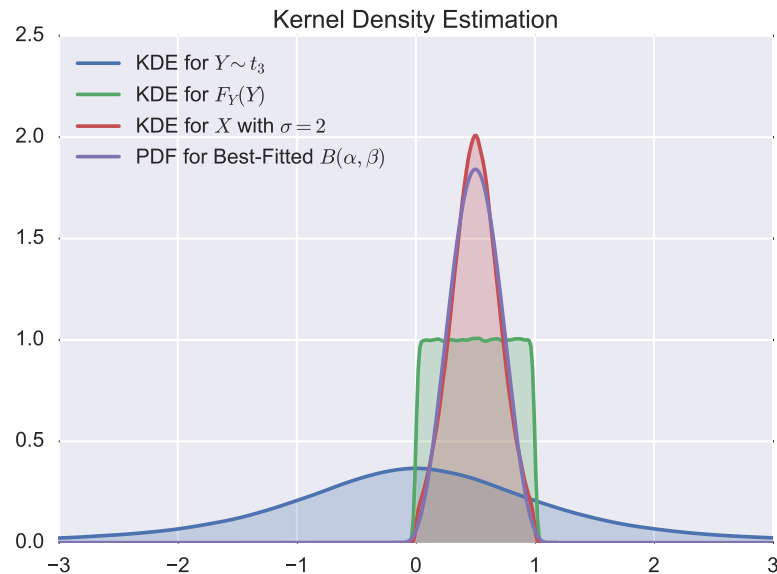


And here's the code responsible for the output above:

```python
### Imports ###
import os
import numpy as np
from numpy import random as rnd
from scipy.stats import t
from scipy.stats import beta
import seaborn as sns
import matplotlib.pyplot as plt

### Parameters ###
nu = 3 # arbitrary choice
sigma = 2 # arbitrary choice

### Simulate Values ###
draws1 = rnd.standard_t(nu, size = 1.e6) # draw one million Y values
draws2 = t.cdf(draws1, df = nu) # classic PIT (F_Y(Y) is standard
    uniform)
draws3 = t.cdf(draws1 / sigma, df = nu) # compute one million X values
    based on Y values
alpha_fit, beta_fit, loc_fit, scale_fit = beta.fit(draws3) # determine
    best-fitted beta

### Plot KDEs and CDF of Best-Fitted Beta PDF ###
kdeFig = plt.figure() # start figure
sns.set("talk") # set seaborn style to "talk" --> increase font size
# add all KDEs with plot labels in LaTeX (hence the use of raw strings
    )
sns.kdeplot(draws1, shade = True, clip = (-3, 3), label = r'KDE for $Y
    \sim t_3$')
```

```
25  sns.kdeplot(draws2, shade = True, clip = (-3, 3), label = r'KDE for
        $F_Y(Y)$')
26  sns.kdeplot(draws3, shade = True, clip = (-3, 3), label = r'KDE for
        $X$ with $\sigma = 2$')
27  # add pdf for best-fitted beta with plot labels in LaTeX (hence the
        use of raw strings)
28  x = np.linspace(-3,3, num = 1000) # create 1000 values between -3 and
        +3
29  y = beta.pdf(x, a = alpha_fit, b = beta_fit, loc = loc_fit, scale =
        scale_fit) # f(x)
30  plt.plot(x, y, label = r'PDF for Best-Fitted $B(\alpha, \beta)$')
31  # title & legend
32  plt.title('Kernel Density Estimation')
33  plt.legend(loc='upper left')
34  os.chdir('../Figures/') # navigate to "Figures" folder
35  plt.savefig('4c.pdf') # store figure as pdf
36  plt.close(kdeFig)
```

(d) Trivially, the answer is: 42

5. (i) The pdf of the Student's $t$ distribution with $n$ degrees of freedom is given by the following expression:

$$f_t(x;n) = \frac{\Gamma(\frac{n+1}{2})n^{\frac{n}{2}}}{\sqrt{\pi}\,\Gamma(\frac{n}{2})}(n+x^2)^{-\frac{n+1}{2}}$$

We're now tasked with finding the cdf of $X \sim t_2$.

**Proposition:** $\qquad F_X(x) = \frac{1}{2}\left(1 + \frac{x}{\sqrt{2+x^2}}\right)$

*Proof.* Let us start by substituting $n = 2$ into the pdf given above.

$$f_X(x;n) = \frac{\Gamma(\frac{n+1}{2})n^{\frac{n}{2}}}{\sqrt{\pi}\,\Gamma(\frac{n}{2})}(n+x^2)^{-\frac{n+1}{2}}$$

$$f_X(x;2) = \frac{\Gamma(\frac{2+1}{2})2^{\frac{2}{2}}}{\sqrt{\pi}\,\Gamma(\frac{2}{2})}(2+x^2)^{-\frac{2+1}{2}}$$

$$= \frac{\Gamma(\frac{3}{2})2}{\sqrt{\pi}\,\Gamma(1)}(2+x^2)^{-\frac{3}{2}}$$

As shown at the beginning of my answer to question 2a), we know the gamma difference equation to be true and can hence rewrite $\Gamma(3/2)$ as $\frac{1}{2}\Gamma(1/2)$. Thus:

$$f_X(x;2) = \frac{\frac{1}{2}\Gamma(\frac{1}{2})\cancel{2}}{\sqrt{\pi}\,\Gamma(1)}(2+x^2)^{-\frac{3}{2}}$$

Additionally, we've shown (question 2b)) that $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(1) = 1$. Using these results allows us to simplify our pdf even more.

$$f_X(x;2) = \frac{\cancel{\sqrt{\pi}}}{\cancel{\sqrt{\pi}}}(2+x^2)^{-\frac{3}{2}}$$

$$= (2+x^2)^{-\frac{3}{2}}$$

$$= \frac{1}{(2+x^2)^{\frac{3}{2}}}$$

In order to obtain the cdf, we now need to integrate our simplified pdf.

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{(2+t^2)^{\frac{3}{2}}} \, dt$$

This looks like a fairly standard trigonometric substitution problem, so let us do the following substitution:

$$t = \sqrt{2}\tan(\theta) \iff \theta = \arctan\left(\frac{t}{\sqrt{2}}\right)$$

Hence:

$$\frac{dt}{d\theta} = \sqrt{2}\sec^2(\theta) \iff dt = \sqrt{2}\sec^2(\theta)\,d\theta$$

Note that substitutions can change the bounds of integration. Since we're planning on undoing our substitutions later on, let us merely use $a$ and $b$ for our (potentially) new limits. Using the proposed substitution and considering the fact that $t^2 = (\sqrt{2}\tan(\theta))^2 = 2\tan^2(\theta)$ yields:

$$F_X(x) = \int_a^b \frac{\sqrt{2}\sec^2(\theta)}{(2+2\tan^2(\theta))^{\frac{3}{2}}} \, d\theta$$

$$= \int_a^b \frac{\sqrt{2}\sec^2(\theta)}{(2\underbrace{(1+\tan^2(\theta)}_{=\sec^2(\theta)})^{\frac{3}{2}}} \, d\theta$$

$$= \int_a^b \frac{\sqrt{2}\sec^2(\theta)}{(2\sec^2(\theta))^{\frac{3}{2}}} \, d\theta$$

$$= \int_a^b \frac{2^{\frac{1}{2}}\sec^2(\theta)}{2^{\frac{3}{2}}\sec^{2\frac{3}{2}}(\theta)} \, d\theta$$

$$= \int_a^b \frac{\sec^2(\theta)}{2\sec^3(\theta)} \, d\theta$$

$$= \frac{1}{2}\int_a^b \frac{1}{\sec(\theta)} \, d\theta$$

$$= \frac{1}{2}\int_a^b \cos(\theta) \, d\theta$$

$$= \frac{1}{2}\sin(\theta)\Big|_a^b$$

Reversing the substitution with $\theta = \arctan(t/\sqrt{2})$ and recalling that $\sin(\arctan(t)) = \frac{t}{\sqrt{1+t^2}}$ produces:

$$F_X(x) = \left[\frac{t}{2^{\frac{3}{2}}\sqrt{\frac{t^2}{2}+1}}\right]_{-\infty}^{x}$$

$$= \left[ \frac{t}{2 \cdot 2^{\frac{1}{2}} \sqrt{\frac{t^2}{2} + 1}} \right]_{-\infty}^{x}$$

$$= \left[ \frac{t}{2\sqrt{t^2 + 2}} \right]_{-\infty}^{x}$$

$$= \frac{x}{2\sqrt{x^2 + 2}} - \underbrace{\lim_{t \to -\infty} \frac{t}{2\sqrt{t^2 + 2}}}_{= -1/2}$$

$$= \frac{x}{2\sqrt{x^2 + 2}} + \frac{1}{2}$$

$$= \frac{1}{2} \left( \frac{x}{\sqrt{x^2 + 2}} + 1 \right)$$

■

Alternatively (and probably somewhat easier), we could have made use of the fundamental theorem of calculus to show this result.

(ii) Deriving the quantile function requires us to invert the cdf, which is assumed to be possible. While in general, there is no closed-form expression for the quantile function of a Student's $t$ distribution, we're able to provide one in this case as $n = 2$. Recall the cdf (given $n = 2$) we derived above and set it equal to $p$.

$$F_X(x) = \frac{1}{2} \left( \frac{x}{\sqrt{x^2 + 2}} + 1 \right) = p$$

We're now interested in rearranging this equation in order to obtain $x$ as a function of $p$ (the quantile function).

$$\frac{1}{2} \left( \frac{x}{\sqrt{x^2 + 2}} + 1 \right) = p$$

$$\iff \frac{x}{\sqrt{x^2 + 2}} + 1 = 2p$$

$$\iff \frac{x}{\sqrt{x^2 + 2}} = 2p - 1$$

$$\iff x = \sqrt{x^2 + 2}(2p - 1)$$

$$\iff x^2 = (x^2 + 2)(2p - 1)^2$$

$$= x^2(2p - 1)^2 + 2(2p - 1)^2$$

$$\iff x^2 - x^2(2p - 1)^2 = 2(2p - 1)^2$$

$$\iff x^2(1 - (2p - 1)^2) = 2(2p - 1)^2$$

$$\iff x^2(1 - (4p^2 - 4p + 1)) = 2(2p - 1)^2$$

$$\iff x^2(1 - 4p^2 + 4p - 1)) = 2(2p - 1)^2$$

$$\iff x^2(4p - 4p^2) = 2(2p - 1)^2$$

$$\iff x^2 4p(1 - p) = 2(2p - 1)^2$$

$$\iff x^2 = \frac{2(2p - 1)^2}{4p(1 - p)}$$

$$= \frac{(2p-1)^2}{2p(1-p)}$$

$$\iff x = \pm\sqrt{\frac{(2p-1)^2}{2p(1-p)}}$$

$$= \pm\frac{2p-1}{\sqrt{2p(1-p)}}$$

Since we know the distribution to be symmetric about zero, we can conclude that the quantile function for $p = 0.5$ should yield zero and negative (positve) values if $p$ is below (above) one half. We can thus immediately discard the negative solution of the two candidates found above. The quantile function is then:

$$F_X^{-1}(p) = \frac{2p-1}{\sqrt{2p(1-p)}}$$

Visually, this quantile function looks as follows and exhibits all desired properties mentioned above:
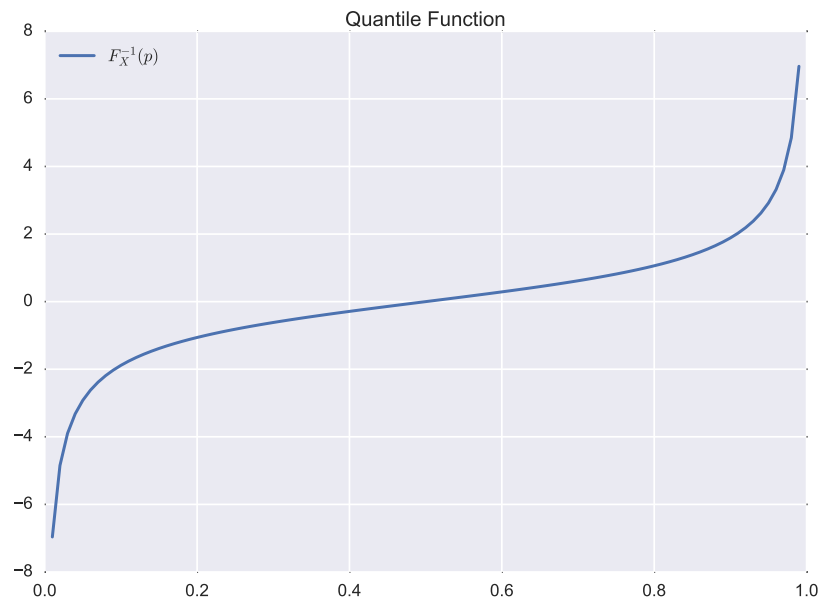


(iii) The second raw moment does not exist in this case.

**Proposition:**     $\mathbb{E}\left[X^2\right]$ is not defined if $n = 2$

*Proof.* In order to show this, let us start with the general case where degrees of freedom are denoted by $n$. Furthermore, we'll use $K$ as defined in the task description statement:

$$K_n = \frac{n^{-\frac{1}{2}}}{B\left(\frac{n}{2}, \frac{1}{2}\right)}$$

Beginning with the definition of the expected value:

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x) \, dx$$

Plugging in the density (using $K$ as defined above) yields:

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 K_n \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \mathrm{d}x$$

Due to symmetry (as given task description statement), we can rewrite the integral as follows:

$$\mathbb{E}\left[X^2\right] = 2\int_0^{\infty} x^2 K_n \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \mathrm{d}x$$

$$= 2K_n \int_0^{\infty} x^2 \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \mathrm{d}x$$

We can now perform a u-substitution with $u = \frac{x^2}{n}$, which implies $x^2 = un$, and $\mathrm{d}x = \frac{\sqrt{n}}{2\sqrt{u}} \mathrm{d}u$. Note that this substitution leaves the bounds of integration unchanged. Hence:

$$\mathbb{E}\left[X^2\right] = 2K_n \int_0^{\infty} un\,(1 + u)^{-\frac{n+1}{2}} \frac{\sqrt{n}}{2\sqrt{u}} \mathrm{d}u$$

$$= \cancel{2}K_n \int_0^{\infty} un\,(1 + u)^{-\frac{n+1}{2}} n^{\frac{1}{2}} u^{-\frac{1}{2}} \frac{\cancel{1}}{\cancel{2}} \mathrm{d}u$$

$$= K_n \int_0^{\infty} u^{\frac{1}{2}} n^{\frac{3}{2}} (1 + u)^{-\frac{n+1}{2}} \mathrm{d}u$$

$$= K_n n^{\frac{3}{2}} \int_0^{\infty} u^{\frac{1}{2}} (1 + u)^{-\frac{n+1}{2}} \mathrm{d}u$$

$$= K_n n^{\frac{3}{2}} \int_0^{\infty} \frac{u^{\frac{1}{2}}}{(1 + u)^{\frac{n+1}{2}}} \mathrm{d}u$$

This integral already looks very similar to the alternate representation of the beta function we've proven for question 2b). Let us thus rewrite the exponents in order to have the exact same form.

$$\mathbb{E}\left[X^2\right] = K_n n^{\frac{3}{2}} \int_0^{\infty} \frac{u^{\frac{3}{2}-1}}{(1 + u)^{\frac{n}{2}+\frac{1}{2}}} \mathrm{d}u$$

$$= K_n n^{\frac{3}{2}} \int_0^{\infty} \frac{u^{\frac{3}{2}-1}}{(1 + u)^{\frac{n}{2}+\frac{1}{2}+\frac{2}{2}-\frac{2}{2}}} \mathrm{d}u$$

$$= K_n n^{\frac{3}{2}} \int_0^{\infty} \frac{u^{\frac{3}{2}-1}}{(1 + u)^{\frac{n}{2}-1+\frac{3}{2}}} \mathrm{d}u$$

If we now define $a = \frac{3}{2}$ and $b = \frac{n}{2} - 1$, we can write:

$$\mathbb{E}\left[X^2\right] = K_n n^{\frac{3}{2}} \int_0^{\infty} \frac{u^{a-1}}{(1 + u)^{b+a}} \mathrm{d}u$$

$$= K_n n^{\frac{3}{2}} \int_0^{\infty} u^{a-1} \left(\frac{1}{u + 1}\right)^{a+b} \mathrm{d}u$$

Now we have the exact same form of the beta function mentined earlier and can thus replace the entire integral with $B\left(\frac{3}{2}, \frac{n}{2} - 1\right)$.

$$\mathbb{E}\left[X^2\right] = K_n n^{\frac{3}{2}} B\left(\frac{3}{2}, \frac{n}{2} - 1\right)$$

Let us now use our definition of $K_n$ since it contains another beta function.

$$\mathbb{E}\left[X^2\right] = \frac{n^{-\frac{1}{2}}}{B\left(\frac{n}{2}, \frac{1}{2}\right)} n^{\frac{3}{2}} B\left(\frac{3}{2}, \frac{n}{2} - 1\right)$$
$$= \frac{n^{\frac{3}{2}}}{n^{\frac{1}{2}}} \frac{B\left(\frac{3}{2}, \frac{n}{2} - 1\right)}{B\left(\frac{n}{2}, \frac{1}{2}\right)}$$
$$= n \frac{B\left(\frac{3}{2}, \frac{n}{2} - 1\right)}{B\left(\frac{n}{2}, \frac{1}{2}\right)}$$

Using the relationship between the beta and gamma function we've proven for question 2c) enables us to write:

$$\mathbb{E}\left[X^2\right] = n \frac{\dfrac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{n}{2} - 1\right)}{\Gamma\left(\frac{3}{2} + \frac{n}{2} - 1\right)}}{\dfrac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2} + \frac{1}{2}\right)}}$$
$$= n \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{n}{2} - 1\right)\Gamma\left(\frac{n}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{3}{2} + \frac{n}{2} - 1\right)\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}$$

Since $\Gamma\left(\frac{3}{2} + \frac{n}{2} - 1\right) = \Gamma\left(\frac{3}{2} + \frac{n}{2} - \frac{2}{2}\right) = \Gamma\left(\frac{1}{2} + \frac{n}{2}\right)$:

$$\mathbb{E}\left[X^2\right] = n \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{n}{2} - 1\right)\cancel{\Gamma\left(\frac{n}{2} + \frac{1}{2}\right)}}{\cancel{\Gamma\left(\frac{1}{2} + \frac{n}{2}\right)}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}$$
$$= n \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{n}{2} - 1\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}$$

Applying the gamma difference equation we've proven at the beginning of question 2a) results in the following simplifications:

$$\mathbb{E}\left[X^2\right] = n \frac{\frac{1}{2}\cancel{\Gamma\left(\frac{1}{2}\right)}\Gamma\left(\frac{n}{2} - 1\right)}{\Gamma\left(\frac{n}{2}\right)\cancel{\Gamma\left(\frac{1}{2}\right)}}$$
$$= n \frac{\Gamma\left(\frac{n}{2} - 1\right)}{2\Gamma\left(\frac{n}{2}\right)}$$
$$= n \frac{\cancel{\Gamma\left(\frac{n}{2} - 1\right)}}{2\left(\frac{n}{2} - 1\right)\cancel{\Gamma\left(\frac{n}{2} - 1\right)}}$$
$$= \frac{n}{2\left(\frac{n}{2} - 1\right)}$$

$$= \frac{n}{n-2}$$

If $n = 2$, this expression is clearly not defined (division by zero) and we've thus shown that the second raw moment does not exist.

■

6. (i) Let us consider a random variable $C \sim \chi_k^2$ with the following density:

$$f_C(c; k) = \frac{1}{2^{k/2}\Gamma(k/2)} c^{k/2-1} e^{-c/2} \mathbb{I}_{(0,\infty)}(c)$$

Letting $X \sim \sqrt{C/k}$ constitutes a simple univariate transformation with $c = kx^2$. Obtaining the pdf of $X$ requires us to use the following formula, which applies if the transformation is strictly monotonic (valid in our case):

$$f_X(x; k) = f_C(c; k) \left| \frac{\mathrm{d}c}{\mathrm{d}x} \right|$$

Let us thus prove this result in order for us to use it during our upcoming derivation. Also note that we've proven a special case of this already (see question 4b)).

**Proposition:** Assuming a random variable $S$ with pdf $f_S(s)$, cdf $F_S(s)$, and a strictly monotonic (required for upcoming inversion) transformation $T = g(S)$, we can express the pdf of $T$ as $f_T(t) = f_S(s) \left| \frac{\mathrm{d}s}{\mathrm{d}t} \right|$ where $f_S(s)$ is to be understood as $f_S(g^{-1}(t))$.

*Proof.* This proof contains two parts (strictly increasing and strictly decreasing transformations). Let us start with the strictly increasing case:

$$\begin{aligned} F_T(t) &= \Pr(T \le t) \\ &= \Pr(g(S) \le t) \\ &= \Pr(g^{-1}(g(S)) \le g^{-1}(t)) \\ &= \Pr(S \le g^{-1}(t)) \\ &= F_S(g^{-1}(t)) \end{aligned}$$

Notice that the direction of the inequality remains unchanged since $g$ is assumed to be a (strictly) increasing function. In order to obtain the pdf we can now take derivatives - using the chain rule - of both sides.

$$\begin{aligned} f_T(t) &= f_S(g^{-1}(t)) \frac{\mathrm{d}g^{-1}(t)}{\mathrm{d}t} \\ &= f_S(g^{-1}(t)) \frac{\mathrm{d}s}{\mathrm{d}t} \\ &= f_S(s) \frac{\mathrm{d}s}{\mathrm{d}t} \end{aligned}$$

And now for the case that $g$ is strictly decreasing:

$$\begin{aligned} F_T(t) &= \Pr(T \le t) \\ &= \Pr(g(S) \le t) \\ &= \Pr(g^{-1}(g(S)) \ge g^{-1}(t)) \\ &= \Pr(S \ge g^{-1}(t)) \\ &= 1 - F_S(g^{-1}(t)) \end{aligned}$$

Note how the inequality was reversed due to $g$ being a (strictly) decreasing function. In order to obtain the pdf we can now, once again, take derivatives - using the chain rule - of both sides.

$$\begin{aligned} f_T(t) &= -f_S(g^{-1}(t)) \frac{\mathrm{d}g^{-1}(t)}{\mathrm{d}t} \\ &= -f_S(g^{-1}(t)) \frac{\mathrm{d}s}{\mathrm{d}t} \\ &= -f_S(s) \frac{\mathrm{d}s}{\mathrm{d}t} \end{aligned}$$

We can now revert our case differentiation and summarize our result. Given a strictly monotonic transformation (irrespective of the fact whether it's strictly increasing or strictly decreasing), we can write the transformed pdf as:

$$f_T(t) = f_S(s) \left| \frac{\mathrm{d}s}{\mathrm{d}t} \right|$$

Recall that $f_S(s)$ is to be understood as $f_S(g^{-1}(t))$.

$$\blacksquare$$

Having shown the validity of the univariate transformation theorem, we're ready to apply it to our problem where $X \sim \sqrt{C/k}$ and $c = kx^2$.

$$\begin{aligned} f_X(x; k) &= f_C(c; k) \left| \frac{\mathrm{d}c}{\mathrm{d}x} \right| \\ &= \frac{1}{2^{k/2}\Gamma(k/2)} (kx^2)^{k/2-1} e^{-(kx^2)/2} \underbrace{2kx}_{=|\mathrm{d}c/\mathrm{d}x|} \mathbb{I}_{(0,\infty)}(kx^2) \end{aligned}$$

Note that $\mathbb{I}_{(0,\infty)}(kx^2)$ is just $\mathbb{I}_{(0,\infty)}(x)$

$$\begin{aligned} &= \frac{2kx}{2^{k/2}\Gamma(k/2)} (kx^2)^{k/2-1} e^{-(kx^2)/2} \mathbb{I}_{(0,\infty)}(x) \\ &= \frac{2kx2^{-k/2}}{\Gamma(k/2)} (kx^2)^{k/2-1} e^{-(kx^2)/2} \mathbb{I}_{(0,\infty)}(x) \\ &= \frac{2^{1-k/2}kx}{\Gamma(k/2)} (kx^2)^{k/2-1} e^{-(kx^2)/2} \mathbb{I}_{(0,\infty)}(x) \end{aligned}$$

$$= \frac{2^{1-k/2}k^{1+k/2-1}x^{1+2k/2-2}}{\Gamma(k/2)} e^{-(kx^2)/2} \, \mathbb{I}_{(0,\infty)}(x)$$

$$= \frac{2^{1-k/2}k^{k/2}x^{k-1}}{\Gamma(k/2)} e^{-(kx^2)/2} \, \mathbb{I}_{(0,\infty)}(x)$$

(ii) Let us use our newly derived pdf and plug in $k = 1$.

$$f_X(x;k) = \frac{2^{1-k/2}k^{k/2}x^{k-1}}{\Gamma(k/2)} e^{-(kx^2)/2} \, \mathbb{I}_{(0,\infty)}(x)$$

$$f_X(x;k) = \frac{2^{1-1/2}1^{1/2}x^{1-1}}{\Gamma(1/2)} e^{-(1x^2)/2} \, \mathbb{I}_{(0,\infty)}(x)$$

$$f_X(x;1) = \frac{2^{1/2}}{\Gamma(1/2)} e^{-x^2/2} \, \mathbb{I}_{(0,\infty)}(x)$$

$$f_X(x;1) = \frac{\sqrt{2}}{\Gamma(1/2)} e^{-x^2/2} \, \mathbb{I}_{(0,\infty)}(x)$$

As shown in my answer to question 1, we know $\Gamma(1/2)$ to be equal to $\sqrt{\pi}$. Thus:

$$f_X(x;k) = \sqrt{\frac{2}{\pi}} e^{-x^2/2} \, \mathbb{I}_{(0,\infty)}(x)$$

This now constitutes the pdf of the folded normal. We can easily compute the mean as follows:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} f_X(x) x \, \mathrm{d}x$$

$$= \int_{-\infty}^{\infty} \sqrt{\frac{2}{\pi}} e^{-x^2/2} \, \mathbb{I}_{(0,\infty)}(x) x \, \mathrm{d}x$$

$$= \int_{0}^{\infty} \sqrt{\frac{2}{\pi}} e^{-x^2/2} x \, \mathrm{d}x$$

$$= \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} e^{-x^2/2} x \, \mathrm{d}x$$

Performing the obvious u-substitution of $u = x^2/2$ leaves the bounds of integration unchanged and implies $\mathrm{d}x = \frac{1}{x} \mathrm{d}u$. Hence:

$$\mathbb{E}[X] = \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} e^{-u} \, \mathrm{d}u$$

$$= \sqrt{\frac{2}{\pi}} \left. -e^{-u} \right|_{0}^{\infty}$$

$$= \sqrt{\frac{2}{\pi}} \left( \lim_{u \to \infty} -e^{-u} - (-e^0) \right)$$

$$= \sqrt{\frac{2}{\pi}} (0 + 1)$$

$$= \sqrt{\frac{2}{\pi}}$$

7. (a) In order to develop the requested closed-form expression for $S = \sum_{j=0}^{u} ja^j$, let us compare $S$ and $aS$ by subtracting $aS$ from $S$, which yields $S(1-a)$. Additionally we'll also compute $aS(1-a)$ and subtract it from $S(1-a)$ resulting in $S(1-a) - aS(1-a)$.

$$
\begin{aligned}
S &= a + 2a^2 + 3a^3 + \ldots + (u-1)a^{u-1} + \quad ua^u \\
aS &= \quad\; a^2 + 2a^3 + \ldots + (u-2)a^{u-1} + (u-1)a^u + \quad ua^{u+1} \\
S(1-a) &= a + \; a^2 + a^3 + \ldots + \quad a^{u-1} \quad + \quad a^u \quad - \quad ua^{u+1} \\
aS(1-a) &= \quad\; a^2 + a^3 + \ldots + \quad a^{u-1} \quad + \quad a^u \quad + \quad a^{u+1} \quad - ua^{u+2} \\
S(1-a) - aS(1-a) &= a \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad -(1+u)a^{u+1} + ua^{u+2}
\end{aligned}
$$

The derivation above now enables us to finalize our result:

$$
\begin{aligned}
S(1-a) - aS(1-a) &= a - (1+u)a^{u+1} + ua^{u+2} \\
\Longleftrightarrow \quad S(1-a)(1-a) &= a - (1+u)a^{u+1} + ua^{u+2} \\
\Longleftrightarrow \quad S(1-a)^2 &= a - (1+u)a^{u+1} + ua^{u+2} \\
\Longleftrightarrow \quad S &= \frac{a - (1+u)a^{u+1} + ua^{u+2}}{(1-a)^2}
\end{aligned}
$$

(b) As given per the task description statement:

$$
\frac{\mathrm{d}}{\mathrm{d}a} \sum_{j=0}^{u} a^{j+1} = \sum_{j=0}^{u} (j+1)a^j \tag{7}
$$

Since our expression $S = \sum_{j=0}^{u} ja^j$ is currently not in the form of either sides of equation (7), let us try to rectify that by transforming it into the right hand side of equation (7). Starting with our initial form of $S$:

$$
S = \sum_{j=0}^{u} ja^j
$$

As the first iteration is zero, we can start our summation at $j = 1$.

$$
S = \sum_{j=1}^{u} ja^j
$$

$$
S = \sum_{j=1}^{u} ja^{j-1}a
$$

$$
S = a \sum_{j=1}^{u} ja^{j-1}
$$

Change of variables with $k = j - 1$, which of course means $j = k + 1$:

$$
S = a \sum_{k=0}^{u-1} (k+1)a^k
$$

Using equation (7):

$$S = a\frac{\mathrm{d}}{\mathrm{d}a}\underbrace{\sum_{j=0}^{u-1}a^{j+1}}_{:=T} \tag{8}$$

Before we're ready to continue along this path, we need to simplify the geometric series $T$. This can be done analogously to the preceding question:

$$
\begin{aligned}
T &= a + a^2 + a^3 + \ldots + a^{u-1} + a^u\\
aT &= \quad\ \ a^2 + a^3 + \ldots + a^{u-1} + a^u + a^{u+1}\\
T(1-a) &= a + \qquad\qquad\qquad\qquad\qquad\quad - a^{u+1}
\end{aligned}
$$

And thus:

$$
\begin{aligned}
T &= \frac{a - a^{u+1}}{1-a}\\[2mm]
&= \frac{a(1 - a^u)}{1-a}
\end{aligned} \tag{9}
$$

Inserting our expression for $T$ from equation (9) into equation (8) yields:

$$
\begin{aligned}
S &= a\frac{\mathrm{d}}{\mathrm{d}a}\frac{a(1-a^u)}{1-a}\\[2mm]
&= a\frac{\mathrm{d}}{\mathrm{d}a}\frac{a - a^{u+1}}{1-a}
\end{aligned}
$$

Using the quotient rule for derivatives:

$$
\begin{aligned}
S &= a\frac{(1-a)(1-(1+u)a^u) - (a - a^{u+1})(-1)}{(1-a)^2}\\[2mm]
S &= a\frac{(1-a)(1-(1+u)a^u) + a - a^{u+1}}{(1-a)^2}\\[2mm]
S &= a\frac{1 - (1+u)a^u - a(1-(1+u)a^u) + a - a^{u+1}}{(1-a)^2}\\[2mm]
S &= a\frac{1 - (1+u)a^u - \cancel{a} + (1+u)a^{u+1} + \cancel{a} - a^{u+1}}{(1-a)^2}\\[2mm]
S &= a\frac{1 - (1+u)a^u + (1+u)a^{u+1} - a^{u+1}}{(1-a)^2}\\[2mm]
S &= a\frac{1 - (1+u)a^u + ua^{u+1}}{(1-a)^2}\\[2mm]
S &= \frac{a - (1+u)a^{u+1} + ua^{u+2}}{(1-a)^2}
\end{aligned}
$$

And now we've obtained the same result as in a).

8. (a) Despite the fact that we're not required to give a formal proof, I've chosen to provide one anyways as it is much easier to digest than the upcoming semi-verbal explanation.

**Proposition:** Assuming $n$ iid continuous random variables $(X_1, X_2, \ldots, X_n)$ with support $-\infty \leq a < b \leq \infty$ and a random variable $Y = \min(X_i)$, the cdf of $Y$ can be expressed as $F_Y(y) = 1 - [1 - F_X(y)]^n$.

*Proof.*

$$F_Y(y) = \Pr(Y \leq y)$$
$$= 1 - \Pr(Y > y)$$
$$= 1 - \Pr(\min(X_i) > y)$$
$$= 1 - \Pr\left(\bigcap_{i=1}^{n}(X_i > y)\right)$$

Due to independence (iid implies independence):

$$F_Y(y) = 1 - \prod_{i=1}^{n} \Pr(X_i > y)$$

Due to identical distributions (iid implies identical distributions):

$$F_Y(y) = 1 - [\Pr(X > y)]^n$$
$$= 1 - [1 - \Pr(X \leq y)]^n$$
$$= 1 - [1 - F_X(y)]^n$$

$$\blacksquare$$

Having proven the proposition, let us quickly recall the setup detailed in the task description statement in order to provide a more verbal explanation.

- $n$ iid continuous random variables: $X_1, X_2, \ldots, X_n$
- support: $-\infty \leq a < b \leq \infty$
- $Y$ has the following definition: $Y = \min(X_i)$

We then claim (and have proven above) that the cdf of $Y$ can be expressed as $F_Y(y) = 1 - [1 - F_X(y)]^n$ for $y \in (a, b)$. Let us now explain why this is the case. $F_Y(y)$ refers to the probability that the random variable $Y$ is less than or equal to a value $y$. According to our definition of $Y$ this is equivalent to the probability that the smallest realization of our $n$ random variables $(X_1, X_2, \ldots, X_n)$ is less than $y$. Since it is not immediately obvious how to compute this, let us use some basic properties from probability theory. We conclude - from the basic axioms of probability - that we can express the probability that $\min(X_i)$ is less to or equal than $y$ as the difference between one (the certain event) and the complement of our probability. Just to ensure we're on the same page (math is so much easier than prose), let us quickly state what we just argued:

$$F_Y(y) = \Pr(Y \leq y) = 1 - \Pr(Y > y)$$

Or equivalently:

$$F_Y(y) = \Pr(\min(X_i) \leq y) = 1 - \Pr(\min(X_i) > y) \tag{10}$$

We now have an expression for our cdf that involves the probability that $\min(X_i)$ is greater than $y$, which is much easier to study than what we had before. The probability that the smallest realization is larger than $y$ is equivalent to the probability that every single realization is larger than $y$. Since we know $X_1, X_2, \ldots, X_n$ to be independent (thanks to them being iid), this probability is equivalent to the product of individual probabilities. We've already used the "independence" aspect of iid, so let us now use the "identically distributed" part. This means that every probability in our product is the same and we can thus use the $n^{\text{th}}$ power of an individual probability as an expression for the probability of the product. Again, let's quickly study the math:

$$\Pr(\min(X_i) > y) = \prod_{i=1}^{n} \Pr(X_i > y)$$
$$= [\Pr(X > y)]^n$$

We can now use the same reasoning from before and express $\Pr(X > y)$ as $1 - \Pr(X \leq y)$. The probability that a random variable $X$ is less than or equal to a value $y$ can - of course - be written as $F_X(y)$. Combining these insights thus tells us that $[\Pr(X > y)]^n$ can be expressed in terms of the cdf of $X$:

$$\Pr(\min(X_i) > y) = [\Pr(X > y)]^n$$
$$= [1 - \Pr(X \leq y)]^n$$
$$= [1 - F_X(y)]^n \tag{11}$$

Recall that we've motivated why one can express the cdf of $Y$ as the difference between one and the probability that the smallest value of $X$ is larger than $y$ (see equation (10)). Now we can just replace this probability with our expression that we derived (semi-verbally) above (see equation (11)) and obtain the given formula. Again, in order to limit potential confusion, here's the math:

$$F_Y(y) = 1 - \Pr(\min(X_i) > y)$$
$$= 1 - [1 - F_X(y)]^n$$

(b) As indicated by the question's wording, we're to use the relationship we've proven and explained in a) in order to derive the cdf of $Z = n \cdot \min(X_i)$. Hence we could just plug in $z/n$ and continue from there, but let's try to motivate the validity of this approach before implementing it. Note that this justification will - obviously - use the very same logic applied in a).

$$F_Z(z) = \Pr(Z \leq z)$$
$$= \Pr(n \cdot \min(X_i) \leq z)$$
$$= \Pr\left(\min(X_i) \leq \frac{z}{n}\right)$$
$$= 1 - \Pr\left(\min(X_i) > \frac{z}{n}\right)$$
$$= 1 - \Pr\left(\bigcap_{i=1}^{n}(X_i > \frac{z}{n})\right)$$

$$\stackrel{iid}{=} 1 - \prod_{i=1}^{n} \Pr\left(X_i > \frac{z}{n}\right)$$

$$\stackrel{iid}{=} 1 - \left[\Pr\left(X > \frac{z}{n}\right)\right]^n$$

$$= 1 - \left[1 - \Pr\left(X \leq \frac{z}{n}\right)\right]^n$$

$$= 1 - \left[1 - F_X\left(\frac{z}{n}\right)\right]^n$$

As mentioned above, we now see that it's indeed possible to plug in $z/n$ into the expression for the cdf derived in a). Let us use the fact that we assume $X \stackrel{iid}{\sim} \mathrm{Unif}(0,1)$ and recall the cdf of a uniform distribution:

$$F_X(x) = \begin{cases} 0 & \text{for} \quad x < a \\ \frac{x-a}{b-a} & \text{for} \quad a \leq x \leq b \\ 1 & \text{for} \quad x > b \end{cases}$$

Since we're dealing with a standard uniform distribution, $a = 0$ and $b = 1$:

$$F_X(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ x & \text{for} \quad 0 \leq x \leq 1 \\ 1 & \text{for} \quad x > 1 \end{cases}$$

Let us now recall our expression for the cdf of $Z$ and use the cdf of $X$ stated in the preceding line. Note the change with respect to the limits for the case differentiation induced by $z/n$.

$$F_Z(z) = 1 - \left[1 - F_X\left(\frac{z}{n}\right)\right]^n$$

$$F_Z(z) = \begin{cases} 1 - [1 - 0]^n & \text{for} \quad z < 0 \\ 1 - \left[1 - \frac{z}{n}\right]^n & \text{for} \quad 0 \leq z \leq n \\ 1 - [1 - 1]^n & \text{for} \quad z > n \end{cases}$$

$$= \begin{cases} 1 - 1^n & \text{for} \quad z < 0 \\ 1 - \left[1 - \frac{z}{n}\right]^n & \text{for} \quad 0 \leq z \leq n \\ 1 - 0^n & \text{for} \quad z > n \end{cases}$$

$$= \begin{cases} 0 & \text{for} \quad z < 0 \\ 1 - \left[1 - \frac{z}{n}\right]^n & \text{for} \quad 0 \leq z \leq n \\ 1 & \text{for} \quad z > n \end{cases}$$

(c) Recall our result from b):

$$F_Z(z) = \begin{cases} 0 & \text{for} \quad z < 0 \\ 1 - \left[1 - \frac{z}{n}\right]^n & \text{for} \quad 0 \leq z \leq n \\ 1 & \text{for} \quad z > n \end{cases}$$

Since this function is obviously differentiable, we can take the derivative (with respect to $z$) - using the chain rule - of both sides of the equation:

$$f_Z(z) = \begin{cases} 0 & \text{for} \quad z < 0 \\ \frac{d}{dz}\left(1 - \left[1 - \frac{z}{n}\right]^n\right) & \text{for} \quad 0 \leq z \leq n \\ 0 & \text{for} \quad z > n \end{cases}$$

$$= \begin{cases} n\left[1 - \frac{z}{n}\right]^{n-1}\frac{1}{n} & \text{for} \quad 0 \leq z \leq n \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \left[1 - \frac{z}{n}\right]^{n-1} & \text{for} \quad 0 \leq z \leq n \\ 0 & \text{otherwise} \end{cases}$$

Having obtained the pdf of $Z$, let us study its behavior as $n \to \infty$:

$$\lim_{n\to\infty} f_Z(z) = \lim_{n\to\infty} \begin{cases} \left[1 - \frac{z}{n}\right]^{n-1} & \text{for} \quad 0 \leq z \leq n \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \lim_{n\to\infty}\left[1 - \frac{z}{n}\right]^{n-1} & \text{for} \quad z \geq 0 \\ 0 & \text{for} \quad z < 0 \end{cases}$$

Note that we can rewrite $\left[1 - \frac{z}{n}\right]^{n-1}$ as $\left(1 - \frac{z}{n}\right)^n \left(1 - \frac{z}{n}\right)^{-1}$ and thus:

$$\lim_{n\to\infty} f_Z(z) = \begin{cases} \lim_{n\to\infty}\left(1 - \frac{z}{n}\right)^n \left(1 - \frac{z}{n}\right)^{-1} & \text{for} \quad z \geq 0 \\ 0 & \text{for} \quad z < 0 \end{cases}$$

Assuming both limits exist:

$$\lim_{n\to\infty} f_Z(z) = \begin{cases} \lim_{n\to\infty}\left(1 - \frac{z}{n}\right)^n \lim_{n\to\infty}\left(1 - \frac{z}{n}\right)^{-1} & \text{for} \quad z \geq 0 \\ 0 & \text{for} \quad z < 0 \end{cases}$$

Trivially, the second limit evaluates to one and thus:

$$\lim_{n\to\infty} f_Z(z) = \begin{cases} \lim_{n\to\infty}\left(1 - \frac{z}{n}\right)^n & \text{for} \quad z \geq 0 \\ 0 & \text{for} \quad z < 0 \end{cases}$$

We recognize the limit representation of $e^{-z}$ and have thus finished our derivation:

$$\lim_{n\to\infty} f_Z(z) = \begin{cases} e^{-z} & \text{for} \quad z \geq 0 \\ 0 & \text{for} \quad z < 0 \end{cases}$$

Or more compactly, using an indicator function:

$$\lim_{n\to\infty} f_Z(z) = e^{-z}\mathbb{I}_{[0,\infty)}(z)$$

We've just shown that this approaches the pdf of an exponential with $\lambda = 1$.

9. Let $X \sim \exp(\lambda)$ and $Y = \exp(X) = e^X$. This implies that $X = \log Y$ (we'll be using log to refer to the natural logarithm). We're now tasked with deriving the pdf of $Y$. Let's apply what we've shown to be true (see question 6i)) with respect to univariate transformations.

$$f_Y(y) = f_X(x) \left| \frac{\mathrm{d}x}{\mathrm{d}y} \right|$$

$$= f_X(\log y) \left| \frac{\mathrm{d}x}{\mathrm{d}y} \right| \tag{12}$$

Recall the pdf of an exponentially distributed random variable $X$.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for} \quad x \geq 0 \\ 0 & \text{for} \quad x < 0 \end{cases}$$

$$= \lambda e^{-\lambda x} \, \mathbb{I}_{[0,\infty)}(x)$$

Inserting this density into equation (12) yields:

$$f_Y(y) = \lambda e^{-\lambda \log y} \, \mathbb{I}_{[0,\infty)}(\log y) \left| \frac{\mathrm{d}x}{\mathrm{d}y} \right|$$

$$= \lambda e^{-\lambda \log y} \, \mathbb{I}_{[0,\infty)}(\log y) \frac{1}{y}$$

$$= \lambda e^{-\lambda \log y} \frac{1}{y} \, \mathbb{I}_{[1,\infty)}(y)$$

$$= \lambda \left( e^{\log y} \right)^{-\lambda} \frac{1}{y} \, \mathbb{I}_{[1,\infty)}(y)$$

$$= \lambda y^{-\lambda} \frac{1}{y} \, \mathbb{I}_{[1,\infty)}(y)$$

$$= \lambda y^{-\lambda-1} \, \mathbb{I}_{[1,\infty)}(y)$$

$$= \lambda y^{-(\lambda+1)} \, \mathbb{I}_{[1,\infty)}(y)$$

Having derived the pdf, we're in a position to compute the $m^{\text{th}}$ moment ($\mathbb{E}[Y^m]$).

$$\mathbb{E}[Y^m] = \int_{-\infty}^{\infty} f_Y(y) y^m \, \mathrm{d}y$$

$$= \int_{-\infty}^{\infty} \lambda y^{-(\lambda+1)} y^m \, \mathbb{I}_{[1,\infty)}(y) \, \mathrm{d}y$$

$$= \int_1^{\infty} \lambda y^{-(\lambda+1)} y^m \, \mathrm{d}y$$

$$= \lambda \int_1^{\infty} y^{m-(\lambda+1)} \, \mathrm{d}y$$

$$= \lambda \int_1^{\infty} y^{m-\lambda-1} \, \mathrm{d}y$$

$$= \lambda \left[ \frac{1}{m-\lambda} y^{m-\lambda} \right]_1^{\infty}$$

$$= \frac{\lambda}{m-\lambda} \left[ y^{m-\lambda} \right]_1^\infty$$

$$= \frac{\lambda}{m-\lambda} \left( \lim_{y\to\infty} y^{m-\lambda} - 1 \right)$$

We now need to differentiate two scenarios. If $m < \lambda$:

$$\mathbb{E}\left[ Y^m \right] = \frac{\lambda}{m-\lambda} \left( \lim_{y\to\infty} y^{m-\lambda} - 1 \right)$$

$$= \frac{\lambda}{m-\lambda} (0 - 1)$$

$$= -\frac{\lambda}{m-\lambda}$$

$$= \frac{\lambda}{\lambda - m}$$

If $m \geq \lambda$, the $m^{\text{th}}$ moment does not exist as the limit $\lim_{y\to\infty} y^{m-\lambda}$ is not finite.

10. Let us consider a random variable $X \sim t_n$ (Student's $t$ distribution with $n$ degrees of freedom) with the following density (see question 5):

$$f_X(x;n) = \frac{\Gamma(\frac{n+1}{2}) n^{\frac{n}{2}}}{\sqrt{\pi}\, \Gamma(\frac{n}{2})} (n + x^2)^{-\frac{n+1}{2}} = \frac{n^{-\frac{1}{2}}}{B\left(\frac{n}{2}, \frac{1}{2}\right)} \left( 1 + \frac{x^2}{n} \right)^{-\frac{n+1}{2}}$$

Since the question's wording already implies a certain approach to tackle the problem, we'll adapt it here and refrain from using the relationship between a Student's t and a Chi-square distribution in conjunction with the continuous mapping theorem, which would have resulted in a (possibly) nicer proof. As the question description statement already mentiones Stirling's approximation, it is used without proof.

**Proposition:** The kernel and constant of integration of a Student's $t$ approach that of a normal:

$$\lim_{n\to\infty} f_X(x;n) = \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}$$

*Proof.* As indicated in the problem description statement, this proof contains two parts. Let us start by showing that the kernel approaches that of a normal. In order to do this, we'll first rearrange the pdf stated above.

$$f_X(x;n) = \frac{n^{-\frac{1}{2}}}{B\left(\frac{n}{2}, \frac{1}{2}\right)} \left( 1 + \frac{x^2}{n} \right)^{-\frac{n+1}{2}}$$

$$= \frac{1}{\sqrt{n} B\left(\frac{n}{2}, \frac{1}{2}\right)} \left( 1 + \frac{x^2}{n} \right)^{-\frac{n+1}{2}}$$

Using the relationship between the beta and gamma function proven in 5c):

$$f_X(x; n) = \frac{1}{\sqrt{n} \dfrac{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2} + \frac{1}{2}\right)}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

$$f_X(x; n) = \frac{\Gamma\left(\frac{n}{2} + \frac{1}{2}\right)}{\sqrt{n}\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{1}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

Recall that we've shown that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ in question 1 and thus:

$$f_X(x; n) = \frac{\Gamma\left(\frac{n}{2} + \frac{1}{2}\right)}{\sqrt{\pi n}\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

$$= \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\Gamma\left(\frac{n}{2}\right)} \underbrace{\left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}}_{\text{Kernel}}$$

Let us now show that this kernel approaches (as $n \to \infty$) a standard normal's kernel, which is equal to $e^{-\frac{x^2}{2}}$.

$$\lim_{n \to \infty} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} = \lim_{n \to \infty} \left(1 + \frac{x^2}{n}\right)^{-\frac{n}{2}} \left(1 + \frac{x^2}{n}\right)^{-\frac{1}{2}}$$

Assuming both limits exist:

$$\lim_{n \to \infty} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} = \lim_{n \to \infty} \left(1 + \frac{x^2}{n}\right)^{-\frac{n}{2}} \underbrace{\lim_{n \to \infty} \left(1 + \frac{x^2}{n}\right)^{-\frac{1}{2}}}_{=1}$$

$$= \lim_{n \to \infty} \left(1 + \frac{x^2}{n}\right)^{-\frac{n}{2}}$$

This already looks fairly close to the limit representation of Euler's number, so let's try to get there.

$$\lim_{n \to \infty} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} = \lim_{n \to \infty} \left(1 + \frac{x^2}{n}\right)^{-\frac{n}{2} \cdot \frac{x^2}{x^2}}$$

$$= \lim_{n \to \infty} \left[\left(1 + \frac{x^2}{n}\right)^{\frac{n}{x^2}}\right]^{-\frac{x^2}{2}}$$

Let us substitute $u = \frac{n}{x^2}$, which implies $x^2 = \frac{n}{u}$, and hence we can use the aforementioned limit representation of Euler's number (note that $n \to \infty$ implies $u \to \infty$):

$$\lim_{n \to \infty} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} = \lim_{u \to \infty} \left[\left(1 + \frac{1}{u}\right)^u\right]^{-\frac{x^2}{2}}$$

$$= e^{-\frac{x^2}{2}}$$

We've now proven that the kernel indeed approaches that of the normal. All that's left to show is that the constant of integration approaches $1/\sqrt{2\pi}$, which is the standard normal's constant of integration. Let us continue by taking the limit of the Student's $t$ constant of integration and applying Stirling's approximation.

$$\lim_{n \to \infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)} \approx \lim_{n \to \infty} \frac{\cancel{\sqrt{2\pi}}\left(\frac{n+1}{2}\right)^{\frac{n}{2}} e^{-\frac{n+1}{2}}}{\sqrt{\pi n}\cancel{\sqrt{2\pi}}\left(\frac{n}{2}\right)^{\frac{n-1}{2}} e^{-\frac{n}{2}}}$$

$$= \lim_{n \to \infty} \frac{\left(\frac{n+1}{2}\right)^{\frac{n}{2}} \cancel{e^{-\frac{n}{2}}} e^{-\frac{1}{2}}}{\sqrt{\pi n}\left(\frac{n}{2}\right)^{\frac{n-1}{2}} \cancel{e^{-\frac{n}{2}}}}$$

$$= \lim_{n \to \infty} \frac{\left(\frac{n+1}{2}\right)^{\frac{n}{2}}}{\sqrt{\pi n}\, e^{\frac{1}{2}}\left(\frac{n}{2}\right)^{\frac{n-1}{2}}}$$

$$= \lim_{n \to \infty} \frac{\left(\frac{n+1}{2}\right)^{\frac{n}{2}}}{\sqrt{\pi n e}\left(\frac{n}{2}\right)^{\frac{n-1}{2}}}$$

$$= \lim_{n \to \infty} \frac{\left(\frac{n+1}{2}\right)^{\frac{n}{2}}}{\sqrt{\pi n e}\left(\frac{n}{2}\right)^{\frac{n}{2}}\left(\frac{n}{2}\right)^{-\frac{1}{2}}}$$

$$= \lim_{n \to \infty} \frac{\left(\frac{n}{2}\right)^{\frac{1}{2}}\left(\frac{n+1}{2}\right)^{\frac{n}{2}}}{\sqrt{\pi n e}\left(\frac{n}{2}\right)^{\frac{n}{2}}}$$

$$= \lim_{n \to \infty} \frac{\left(\frac{n+1}{2}\right)^{\frac{n}{2}}}{\sqrt{2\pi e}\left(\frac{n}{2}\right)^{\frac{n}{2}}}$$

$$= \lim_{n \to \infty} \frac{1}{\sqrt{2\pi e}}\left(\frac{\frac{n+1}{\cancel{2}}}{\frac{n}{\cancel{2}}}\right)^{\frac{n}{2}}$$

$$= \lim_{n \to \infty} \frac{1}{\sqrt{2\pi e}}\left(\frac{n+1}{n}\right)^{\frac{n}{2}}$$

$$= \frac{1}{\sqrt{2\pi e}} \lim_{n \to \infty}\left(\frac{n+1}{n}\right)^{\frac{n}{2}}$$

$$= \frac{1}{\sqrt{2\pi e}} \lim_{n \to \infty} \sqrt{\left(\frac{n+1}{n}\right)^n}$$

$$= \frac{1}{\sqrt{2\pi e}} \lim_{n \to \infty} \sqrt{\left(1 + \frac{1}{n}\right)^n}$$

Using the limit representation of Euler's number once again:

$$\lim_{n \to \infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\, \Gamma\left(\frac{n}{2}\right)} \approx \frac{1}{\sqrt{2\pi e}} \sqrt{e}$$

$$= \frac{1}{\sqrt{2\pi}}$$

Now we've also shown that the constant of integration approaches that of a standard normal and hence proven the entire proposition.

∎

11. **Proposition:** The finiteness of $\mathbb{E}[X]$ is equivalent to the finiteness of $\mathbb{E}[|X|]$

$$|\mathbb{E}[X]| < \infty \iff \mathbb{E}[|X|] < \infty$$

*Proof.* Let us state the following two definitions

$$X^+ = X\mathbb{I}(X \geq 0)$$
$$X^- = -X\mathbb{I}(X < 0)$$

This means that both $X^+$ and $X^-$ are non-negative and enables us to express $X$ and $|X|$ as follows:

$$X = X^+ - X^- \tag{13}$$
$$|X| = X^+ + X^- \tag{14}$$

Proving the proposition requires us to show two implications, as the proposition contains an equivalence relation. Let us start by proving the following direction first, as it is easier:

$$\mathbb{E}[|X|] < \infty \Rightarrow |\mathbb{E}[X]| < \infty$$

Beginning with our expression for $X$ given by equation (13):

$$X = X^+ - X^-$$
$$\iff \mathbb{E}[X] = \mathbb{E}[X^+ - X^-]$$
$$= \mathbb{E}[X^+] - \mathbb{E}[X^-]$$

Taking the absolute value of both sides of the equation produces:

$$|\mathbb{E}\left[X\right]| = |\mathbb{E}\left[X^+\right] - \mathbb{E}\left[X^-\right]|$$

Applying the famous triangle inequality ($|a - b| \leq |a| + |b|$) yields:

$$|\mathbb{E}\left[X\right]| \leq |\mathbb{E}\left[X^+\right]| + |\mathbb{E}\left[X^-\right]|$$

As mentioned above, $X^+$ and $X^-$ are non-negative and thus - by the definition of the expected value - $\mathbb{E}\left[X^+\right]$ and $\mathbb{E}\left[X^-\right]$ are non-negative as well. Hence, we're allowed to omit the absolute value signs around the expected values on the right hand side of the equation:

$$\begin{aligned}|\mathbb{E}\left[X\right]| &\leq \mathbb{E}\left[X^+\right] + \mathbb{E}\left[X^-\right] \\ &= \mathbb{E}\left[X^+ + X^-\right] \\ &= \mathbb{E}\left[|X|\right]\end{aligned}$$

We have thus shown:

$$|\mathbb{E}\left[X\right]| \leq \mathbb{E}\left[|X|\right]$$

If $\mathbb{E}\left[|X|\right] < \infty$, $|\mathbb{E}\left[X\right]|$ must be finite as well, as it is less than or equal to $\mathbb{E}\left[|X|\right]$.

$$\mathbb{E}\left[|X|\right] < \infty \Rightarrow |\mathbb{E}\left[X\right]| < \infty$$

We have thus proven one of the two implications and are ready to move on to the trickier part of this proof. Let us now show the following implication:

$$|\mathbb{E}\left[X\right]| < \infty \Rightarrow \mathbb{E}\left[|X|\right] < \infty$$

Let's start by assuming that $|\mathbb{E}\left[X\right]| < \infty$ is indeed true and attempt to show that this indeed implies the desired result. In order to simplify our notation and enhance readability, let us denote the statement $|\mathbb{E}\left[X\right]| < \infty$ as $A$.

$$A \Rightarrow -\infty < \mathbb{E}\left[X\right] < \infty$$

Using the expression for $X$ given by equation (13):

$$\begin{aligned}&A \Rightarrow -\infty < \mathbb{E}\left[X^+ - X^-\right] < \infty \\ \Longleftrightarrow\ &A \Rightarrow -\infty < \mathbb{E}\left[X^+\right] - \mathbb{E}\left[X^-\right] < \infty \\ \Longleftrightarrow\ &A \Rightarrow -\infty < \mathbb{E}\left[X^+\right] + \mathbb{E}\left[-X^-\right] < \infty\end{aligned}$$

Given our assumption that $|\mathbb{E}\left[X\right]| < \infty$ is true, we can obviosuly conclude:

$$A \Rightarrow \left(-\infty < \mathbb{E}\left[X^+\right] < \infty \wedge -\infty < \mathbb{E}\left[-X^-\right] < \infty\right)$$

If $-\infty < \mathbb{E}\left[-X^-\right] < \infty$ holds true, so does $-\infty < \mathbb{E}\left[X^-\right] < \infty$ and hence:

$$A \Rightarrow \left(-\infty < \mathbb{E}\left[X^+\right] < \infty \wedge -\infty < \mathbb{E}\left[X^-\right] < \infty\right)$$

Since $\mathbb{E}\left[X^+\right] < \infty$ and $\mathbb{E}\left[X^-\right] < \infty$, we can reverse our separation and deduce that their sum is finite as well:

$$A \Rightarrow \mathbb{E}\left[X^+\right] + \mathbb{E}\left[X^-\right] < \infty$$
$$\Longleftrightarrow \quad A \Rightarrow \mathbb{E}\left[X^+ + X^-\right] < \infty$$

Using our expression for $|X|$ from equation (14):

$$A \Rightarrow \mathbb{E}\left[|X|\right] < \infty$$

We can now recall our definition of $A$ and realize that we've proven the second implication and thus the entire proposition.

$$|\mathbb{E}\left[X\right]| < \infty \Rightarrow \mathbb{E}\left[|X|\right] < \infty$$

$$\blacksquare$$

12. In the context of lower partial moments, we require an expression for

$$T_{h,c}(X) = \int_{-\infty}^{c} x^h f_X(x) \; \mathrm{d}x \tag{15}$$

Let $Z \sim \mathrm{N}(0,1)$ and assume $c < 0$ and $h \in \mathbb{N}$. We're now asked to prove the following proposition.

**Proposition:**
$$T_{h,c}(Z) = \frac{(-1)^h 2^{h/2-1}}{\sqrt{\pi}} \left[ \Gamma\left(\frac{h+1}{2}\right) - \Gamma_{c^2/2}\left(\frac{h+1}{2}\right) \right]$$

*Proof.* Since $Z$ is (standard) normally distributed, let us recall the pdf of a standard normal:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Plugging this density into equation (15)

$$T_{h,c}(Z) = \int_{-\infty}^{c} z^h f_Z(z) \; \mathrm{d}z$$
$$= \int_{-\infty}^{c} z^h \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \; \mathrm{d}z$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} z^h e^{-\frac{z^2}{2}} \; \mathrm{d}z$$

Since $c$ is assumed to be less than zero, we can rewrite our integral as follows:

$$T_{h,c}(Z) = \frac{1}{\sqrt{2\pi}} \left[ \int_{-\infty}^{0} z^h e^{-\frac{z^2}{2}} \; \mathrm{d}z - \int_{c}^{0} z^h e^{-\frac{z^2}{2}} \; \mathrm{d}z \right]$$
$$= \frac{1}{\sqrt{2\pi}} \left[ \int_{-\infty}^{0} z^h e^{-\frac{z^2}{2}} \; \mathrm{d}z + \int_{0}^{c} z^h e^{-\frac{z^2}{2}} \; \mathrm{d}z \right]$$

Due to the symmetric nature of a normal's pdf:

$$T_{h,c}(Z) = \frac{1}{\sqrt{2\pi}}\left[-\int_0^\infty z^h e^{-\frac{z^2}{2}}\,\mathrm{d}z + \int_0^c z^h e^{-\frac{z^2}{2}}\,\mathrm{d}z\right]$$

We'll now perform a u-substitution with $u = \frac{z^2}{2}$, which implies, assuming (wlg) a negative value for z, $z = (-1)(2u)^{1/2}$. Furthermore: $\mathrm{d}z = (-1)(2u)^{-1/2}\,\mathrm{d}u$. Using said substitution then yields:

$$T_{h,c}(Z) = \frac{1}{\sqrt{2\pi}}\left[\cancel{-}\int_0^\infty \left((-1)(2u)^{1/2}\right)^h e^{-u}\cancel{(-1)}(2u)^{-1/2}\,\mathrm{d}u\right.$$
$$\left. + \int_0^{\frac{c^2}{2}} \left((-1)(2u)^{1/2}\right)^h e^{-u}(-1)(2u)^{-1/2}\,\mathrm{d}u\right]$$
$$= \frac{1}{\sqrt{2\pi}}\left[\int_0^\infty (-1)^h(2u)^{h/2}e^{-u}(2u)^{-1/2}\,\mathrm{d}u - \int_0^{\frac{c^2}{2}} (-1)^h(2u)^{h/2}e^{-u}(2u)^{-1/2}\,\mathrm{d}u\right]$$

Combining the various $2u$ terms simplifies our expression even further:

$$T_{h,c}(Z) = \frac{1}{\sqrt{2\pi}}\left[\int_0^\infty (-1)^h(2u)^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u - \int_0^{\frac{c^2}{2}} (-1)^h(2u)^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u\right]$$
$$= \frac{1}{\sqrt{2\pi}}\left[(-1)^h\int_0^\infty (2u)^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u - (-1)^h\int_0^{\frac{c^2}{2}} (2u)^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u\right]$$
$$= \frac{1}{\sqrt{2\pi}}\left[(-1)^h\int_0^\infty 2^{\frac{h-1}{2}}u^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u - (-1)^h\int_0^{\frac{c^2}{2}} 2^{\frac{h-1}{2}}u^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u\right]$$
$$= \frac{1}{\sqrt{2\pi}}\left[(-1)^h 2^{\frac{h-1}{2}}\int_0^\infty u^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u - (-1)^h 2^{\frac{h-1}{2}}\int_0^{\frac{c^2}{2}} u^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u\right]$$
$$= \frac{1}{\sqrt{2\pi}}\left[(-1)^h 2^{\frac{h-1}{2}}\left(\int_0^\infty u^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u - \int_0^{\frac{c^2}{2}} u^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u\right)\right]$$
$$= \frac{(-1)^h 2^{\frac{h-1}{2}}}{\sqrt{2\pi}}\left[\int_0^\infty u^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u - \int_0^{\frac{c^2}{2}} u^{\frac{h-1}{2}}e^{-u}\,\mathrm{d}u\right]$$

We're now almost able to detect a gamma function and an incomplete gamma function. In order to finalize our proof, realize that we can express $\frac{h-1}{2}$ as follows:

$$\frac{h-1}{2} = \frac{h+1}{2} - \frac{2}{2} = \frac{h+1}{2} - 1$$

Hence:

$$T_{h,c}(Z) = \frac{(-1)^h 2^{\frac{h-1}{2}}}{\sqrt{2\pi}}\left[\int_0^\infty u^{\frac{h+1}{2}-1}e^{-u}\,\mathrm{d}u - \int_0^{\frac{c^2}{2}} u^{\frac{h+1}{2}-1}e^{-u}\,\mathrm{d}u\right]$$

Applying the integral definitions of the gamma function and the incomplete gamma function yields the final result:

$$T_{h,c}(Z) = \frac{(-1)^h \, 2^{h/2-1}}{\sqrt{\pi}} \left[ \Gamma\left(\frac{h+1}{2}\right) - \Gamma_{c^2/2}\left(\frac{h+1}{2}\right) \right]$$

■

13. For $m > 2$, let $B \sim \text{Beta}((m-1)/2, (m-1)/2)$ independent of $X \sim \chi_m^2$ with the following pdf:

$$f_X(x; m) = \frac{1}{2^{m/2}\Gamma(m/2)} x^{m/2-1} e^{-x/2} \, \mathbb{I}_{(0,\infty)}(x).$$

Let us further define the following three random variables:

$$S = 2B - 1$$
$$Y = \sqrt{X}$$
$$P = SY$$

In order to derive and integral expression for the density of $P$, we first need to compute the pdfs of $S$ and $Y$, as the product of these two random variables defines $P$. Luckily, obtaining these two pdfs is a fairly easy task as we can apply the simple univariate transformation formula proven in question 6 (a). Let us start by deriving $f_S(s)$.

$$f_S(s; m) = f_B(b; m) \left| \frac{db}{ds} \right|$$

Note that we can express $s$ as $s = 2b - 1$ and $b$ as $b = \frac{s+1}{2}$. This also implies that $\frac{db}{ds} = \frac{1}{2}$. Hence:

$$f_S(s; m) = f_B\left(\frac{s+1}{2}; m\right) \left| \frac{1}{2} \right|$$
$$= f_B\left(\frac{s+1}{2}; m\right) \frac{1}{2} \tag{16}$$

Since $B$ follows a beta distribution with $\alpha = \beta = (m-1)/2$, let us state the corresponding pdf:

$$f_B(b; m) = \frac{b^{\frac{m-1}{2}-1}(1-b)^{\frac{m-1}{2}-1}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \mathbb{I}_{(0,1)}(b)$$

Recalling that $b = \frac{s+1}{2}$:

$$f_B\left(\frac{s+1}{2}; m\right) = \frac{\left(\frac{s+1}{2}\right)^{\frac{m-1}{2}-1}\left(1 - \frac{s+1}{2}\right)^{\frac{m-1}{2}-1}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \mathbb{I}_{(0,1)}\left(\frac{s+1}{2}\right)$$
$$= \frac{\left(\frac{s+1}{2}\right)^{\frac{m-1}{2}-1}\left(1 - \frac{s+1}{2}\right)^{\frac{m-1}{2}-1}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \mathbb{I}_{(-1,1)}(s)$$

$$= \frac{\left(\frac{s+1}{2}\right)^{\frac{m-3}{2}} \left(1 - \frac{s+1}{2}\right)^{\frac{m-3}{2}}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \mathbb{I}_{(-1,1)}(s)$$

Let us now plug this pdf into equation (16).

$$f_S(s; m) = f_B\left(\frac{s+1}{2}; m\right) \frac{1}{2}$$

$$= \frac{1}{2} \frac{\left(\frac{s+1}{2}\right)^{\frac{m-3}{2}} \left(1 - \frac{s+1}{2}\right)^{\frac{m-3}{2}}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \mathbb{I}_{(-1,1)}(s)$$

$$= \frac{\frac{1}{2} \left(\frac{1}{2}\right)^{\frac{m-3}{2}} (s+1)^{\frac{m-3}{2}} \left(\frac{1}{2}\right)^{\frac{m-3}{2}} (2 - (s+1))^{\frac{m-3}{2}}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \mathbb{I}_{(-1,1)}(s)$$

$$= \frac{2^{2-m} (s+1)^{\frac{m-3}{2}} (2 - (s+1))^{\frac{m-3}{2}}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \mathbb{I}_{(-1,1)}(s)$$

$$= \frac{2^{2-m} ((s+1)(2 - s - 1))^{\frac{m-3}{2}}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \mathbb{I}_{(-1,1)}(s)$$

$$= \frac{2^{2-m} (2s - s^2 - s + 2 - s - 1)^{\frac{m-3}{2}}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \mathbb{I}_{(-1,1)}(s)$$

$$= \frac{2^{2-m}}{B\left(\frac{m-1}{2}, \frac{m-1}{2}\right)} \left(1 - s^2\right)^{\frac{m-3}{2}} \mathbb{I}_{(-1,1)}(s)$$

Having derived the density of $S$, let us compute the density of $Y$.

$$f_Y(y; m) = f_X(x; m) \left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|$$

Note that we can express $y$ as $y = \sqrt{x}$ and $x$ as $x = y^2$. This also implies that $\frac{\mathrm{d}x}{\mathrm{d}y} = 2y$. Hence:

$$f_Y(y; m) = f_X(y^2; m) 2y$$

Using the pdf of $X$ we stated at the beginning and recalling that $x = y^2$:

$$f_Y(y; m) = 2y \frac{1}{2^{m/2}\Gamma(m/2)} (y^2)^{m/2-1} e^{-y^2/2} \mathbb{I}_{(0,\infty)}(y^2)$$

$$= \frac{2y}{2^{m/2}\Gamma(m/2)} y^{2m/2-2} e^{-y^2/2} \mathbb{I}_{(0,\infty)}(y)$$

$$= \frac{2y}{2^{m/2}\Gamma(m/2)} y^{m-2} e^{-y^2/2} \mathbb{I}_{(0,\infty)}(y)$$

$$= \frac{2}{2^{m/2}\Gamma(m/2)} y^{m-1} e^{-y^2/2} \mathbb{I}_{(0,\infty)}(y)$$

$$= \frac{2^{1-m/2}}{\Gamma(m/2)} y^{m-1} e^{-y^2/2} \mathbb{I}_{(0,\infty)}(y)$$

We're now in a position to derive the density of $P$. This is slightly trickier than the two derivations we just did as we're now dealing with a product ($P = YS$) of two random variables. In order to use a bivariate transformation, a second random variable $T$ - acting as a dummy - is required, since $P$ is univariate and not bivariate. By carefully choosing our dummy (e.g. $T = S$), we create a bivariate structure and can thus perform a bivariate transformation. Note that $(p, t) = (g_1(y, s), g_2(y, s)) = (ys, s)$ is a bijection where the inverse transform is $(y, s) = (g_1^{-1}(p, t), g_2^{-1}(p, t)) = (\frac{p}{t}, t)$. It is now possible to state the joint density of $P$ and $T$ as follows:

$$f_{P,T}(p, t; m) = f_{Y,S}(y, s; m)|\det \mathbf{J}|$$

Here, $|\det \mathbf{J}|$ refers to the absolute value of the determinant of the Jacobian matrix:

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial y}{\partial p} & \dfrac{\partial y}{\partial t} \\ \dfrac{\partial s}{\partial p} & \dfrac{\partial s}{\partial t} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{t} & -pt^{-2} \\ 0 & 1 \end{bmatrix}$$

$$|\det \mathbf{J}| = \left| \left( \frac{1}{t} \cdot 1 - 0 \cdot (-pt^{-2}) \right) \right| = \left| \frac{1}{t} \right|$$

Inserting this into our joint density yields:

$$f_{P,T}(p, t; m) = f_{Y,S}(y, s; m) \left| \frac{1}{t} \right|$$
$$= f_{Y,S} \left( \frac{p}{t}, t; m \right) |t|^{-1}$$

While we now have an expression for the joint density of $P$ and our dummy $T$, we're only interested in the marginal density of $P$. Let us thus integrate our joint density over $t$.

$$f_P(p; m) = \int_{-\infty}^{\infty} |t|^{-1} f_{Y,S} \left( \frac{p}{t}, t; m \right) \, dt$$

Given that the joint pdf of $Y$ and $S$ assigns densities to pairs of $Y$ and $S$, only the intersection of their respective support is of any consequence. Hence, we can adjust the integral's bounds as follows:

$$f_P(p; m) = \int_0^1 \frac{1}{t} f_{Y,S} \left( \frac{p}{t}, t; m \right) \, dt$$

Note that the new bounds of integration allow us to drop the absolute value on $t$ as it will never be negative. Since $X$ and $B$ are independent, $S = 2B - 1$ and $Y = \sqrt{X}$ are independent as well, which means we can express their joint density as the product of their marginals:

$$f_P(p; m) = \int_0^1 \frac{1}{t} f_Y \left( \frac{p}{t}; m \right) f_S(t; m) \, dt$$

We can now insert the pdfs of $Y$ and $S$ without the indicator functions as we've already adjusted the bounds of integration.

$$f_P(p;m) = \int_0^1 \underbrace{\frac{1}{t} \frac{2^{1-m/2}}{\Gamma\left(\frac{m}{2}\right)} \left(\frac{p}{t}\right)^{m-1} e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2}}_{f_Y(p/t;m)} \underbrace{\frac{2^{2-m}}{B\left(\frac{m-1}{2},\frac{m-1}{2}\right)} \left(1-t^2\right)^{\frac{m-3}{2}}}_{f_S(t;m)} \, dt$$

$$= \frac{2^{1-m/2}}{\Gamma\left(\frac{m}{2}\right)} \frac{2^{2-m}}{B\left(\frac{m-1}{2},\frac{m-1}{2}\right)} \int_0^1 \frac{1}{t} \left(\frac{p}{t}\right)^{m-1} e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \left(1-t^2\right)^{\frac{m-3}{2}} \, dt$$

$$= \frac{2^{1-m/2+2-m}}{\Gamma\left(\frac{m}{2}\right) B\left(\frac{m-1}{2},\frac{m-1}{2}\right)} \int_0^1 \frac{1}{t} \left(\frac{p}{t}\right)^{m-1} e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \left(1-t^2\right)^{\frac{m-3}{2}} \, dt$$

$$= \frac{2^{3(1-m/2)}}{\Gamma\left(\frac{m}{2}\right) B\left(\frac{m-1}{2},\frac{m-1}{2}\right)} \int_0^1 \frac{1}{t} \left(\frac{p}{t}\right)^{m-1} e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \left(1-t^2\right)^{\frac{m-3}{2}} \, dt$$

As we already have a gamma function in our expression, let us convert the beta function using the relationship proven in question 2(c):

$$f_P(p;m) = \frac{2^{3(1-m/2)}\Gamma\left(\frac{m-1}{2}+\frac{m-1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{m-1}{2}\right)\Gamma\left(\frac{m-1}{2}\right)} \int_0^1 \frac{1}{t} \left(\frac{p}{t}\right)^{m-1} e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \left(1-t^2\right)^{\frac{m-3}{2}} \, dt$$

$$= \frac{2^{3(1-m/2)}\Gamma\left(m-1\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{m-1}{2}\right)\Gamma\left(\frac{m-1}{2}\right)} \int_0^1 \frac{1}{t} \left(\frac{p}{t}\right)^{m-1} e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \left(1-t^2\right)^{\frac{m-3}{2}} \, dt$$

Since we now have a general expression for the pdf of $P$, let us prove that in the case of $m=3$, $P$ follows a standard normal distribution ($P \sim \mathrm{N}(0,1)$).

**Proposition:** If $m=3$, $P$ is standard normally distributed and thus has the folowing density:

$$f_P(p;3) = \frac{1}{\sqrt{2\pi}} e^{-\frac{p^2}{2}}$$

*Proof.*

$$f_P(p;3) = \frac{2^{3(1-3/2)}\Gamma\left(3-1\right)}{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{3-1}{2}\right)\Gamma\left(\frac{3-1}{2}\right)} \int_0^1 \frac{1}{t} \left(\frac{p}{t}\right)^{3-1} e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \left(1-t^2\right)^{\frac{3-3}{2}} \, dt$$

$$= \frac{2^{-3/2}\Gamma\left(2\right)}{\Gamma\left(\frac{3}{2}\right)\Gamma\left(1\right)\Gamma\left(1\right)} \int_0^1 \frac{1}{t} \left(\frac{p}{t}\right)^2 e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \, dt$$

Recall - as shown in question 1(b) - that $\Gamma(1)=1$:

$$f_P(p;3) = \frac{2^{-3/2}\Gamma\left(2\right)}{\Gamma\left(\frac{3}{2}\right)} \int_0^1 \frac{1}{t} \left(\frac{p}{t}\right)^2 e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \, dt$$

Applying the gamma difference equation proven in 2(b) then yields:

$$f_P(p;3) = \frac{2^{-3/2} 1 \cdot \cancel{\Gamma\left(1\right)}}{\frac{1}{2}\Gamma\left(\frac{1}{2}\right)} \int_0^1 \frac{1}{t} \left(\frac{p}{t}\right)^2 e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \, dt$$

$$= \frac{1}{2^{3/2}2^{-1}\Gamma\left(\frac{1}{2}\right)} \int_0^1 \frac{1}{t}\left(\frac{p}{t}\right)^2 e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \, dt$$

$$= \frac{1}{\sqrt{2}\Gamma\left(\frac{1}{2}\right)} \int_0^1 \frac{1}{t}\left(\frac{p}{t}\right)^2 e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \, dt$$

Using the fact that $\Gamma(1/2) = \sqrt{\pi}$ as shown in question 1:

$$f_P(p;3) = \frac{1}{\sqrt{2\pi}} \int_0^1 \frac{1}{t}\left(\frac{p}{t}\right)^2 e^{-\frac{1}{2}\left(\frac{p}{t}\right)^2} \, dt$$

Let us now perform a u-substituion with $u = -\frac{1}{2}\left(\frac{p}{t}\right)^2$, which implies $du = \frac{1}{t}\left(\frac{p}{t}\right)^2 dt$ and hence:

$$f_P(p;3) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-p^2/2} e^u \, du$$

$$= \frac{1}{\sqrt{2\pi}} \left[e^u\right]_{-\infty}^{-p^2/2}$$

$$= \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{p^2}{2}} - \lim_{u\to-\infty} e^u\right)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{p^2}{2}}$$

$$\blacksquare$$

We're now tasked to show (using numeric integration) that for every $m > 2$, $P$ also follows a standard normal distribution. I've chosen to demonstrate this with a small Python script using SciPy as an interface to the Fortran library $QUADPACK$. Using adaptive quadrature we first approximate the integral expression of the general form of the pdf and then integrate (again using adaptive quadrature) over the pdf (from $-\infty$ to $p$, where $p$ denotes one of 50 reference values between $-3$ and 3). This procedure is done for every other value of $m > 2$ until we've reached 9 (arbitrary limit to keep processing time reasonable). Using these values we're able to approximate four cdfs (one for each $m$). We then plot all these functions in conjunction with the real cdf of a standard normal. As evidenced by the figure below, we can't distinguish them (as we expected / hoped):

CDFs Approximated Using Numeric Integration)
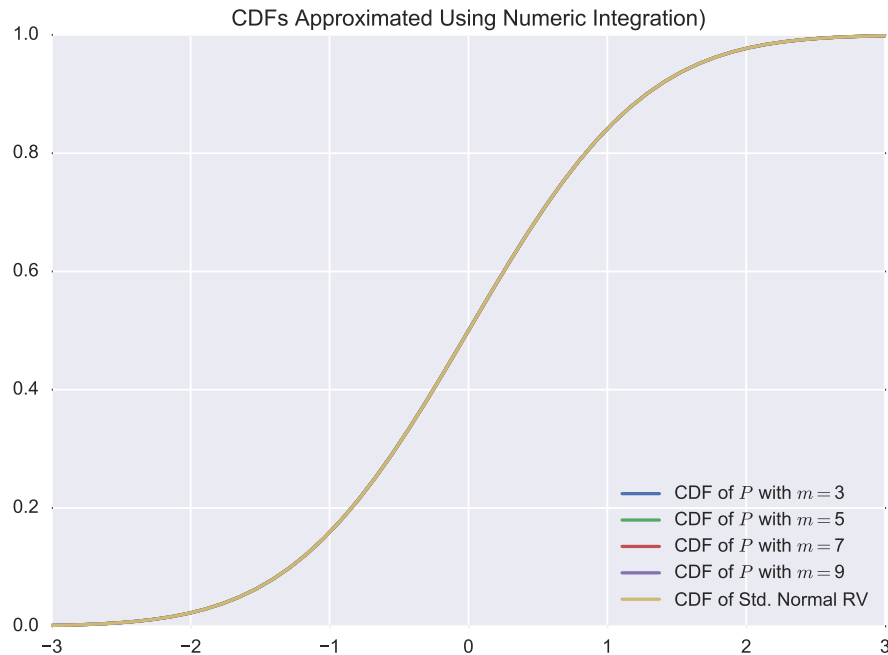
Additionally, I've chosen to store the last iteration's ($m = 9$) approximations and - using the Python module *Pandas* - stored them in a dataframe, added a column containing the values from a normal cdf, and extended it with the absolute value of the difference as an additional way to convinve you of the results. This dataframe is converted to LaTeX code and looks as follows (note that every other row was dropped to keep it at a reasonable size):

| Input Value | Num. Integration | Normal CDF | Absolute Difference |
| --- | --- | --- | --- |
| -3.000000 | 0.001350 | 0.001350 | 2.039601e-15 |
| -2.755102 | 0.002934 | 0.002934 | 2.346214e-15 |
| -2.510204 | 0.006033 | 0.006033 | 2.076464e-15 |
| -2.265306 | 0.011747 | 0.011747 | 7.372575e-16 |
| -2.020408 | 0.021671 | 0.021671 | 2.064286e-13 |
| -1.775510 | 0.037907 | 0.037907 | 8.556350e-14 |
| -1.530612 | 0.062933 | 0.062933 | 9.714451e-17 |
| -1.285714 | 0.099271 | 0.099271 | 7.910339e-16 |
| -1.040816 | 0.148980 | 0.148980 | 1.387779e-16 |
| -0.795918 | 0.213040 | 0.213040 | 5.828671e-16 |
| -0.551020 | 0.290810 | 0.290810 | 1.387779e-15 |
| -0.306122 | 0.379756 | 0.379756 | 1.110223e-15 |
| -0.061224 | 0.475590 | 0.475590 | 1.110223e-16 |
| 0.183673 | 0.572865 | 0.572865 | 9.992007e-16 |
| 0.428571 | 0.665882 | 0.665882 | 4.418688e-14 |
| 0.673469 | 0.749676 | 0.749676 | 2.712275e-13 |
| 0.918367 | 0.820787 | 0.820787 | 2.220446e-16 |
| 1.163265 | 0.877639 | 0.877639 | 3.108624e-15 |
| 1.408163 | 0.920459 | 0.920459 | 1.299383e-11 |
| 1.653061 | 0.950841 | 0.950841 | 1.221245e-15 |

| | | | |
|---|---|---|---|
| 1.897959 | 0.971149 | 0.971149 | 2.153833e-14 |
| 2.142857 | 0.983938 | 0.983938 | 8.881784e-16 |
| 2.387755 | 0.991524 | 0.991524 | 1.043610e-14 |
| 2.632653 | 0.995764 | 0.995764 | 1.110223e-16 |
| 2.877551 | 0.997996 | 0.997996 | 1.798561e-14 |
| 3.000000 | 0.998650 | 0.998650 | 6.328271e-14 |

As indicated by this table, the differences are extremely small and thus, for all intents and purposes, we've shown what was to be shown. Finally, here's the code responsible for the entire analysis:

```python
### Imports ###
import math
import numpy as np
import scipy.special as special
from scipy.integrate import quad, dblquad
from scipy.stats import norm
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
import os
import codecs

### Define & Create Variables ###
pList = np.linspace(-3, 3, num=50) # generate 50 values for p for
    comparison
mList = range(3,10,2) # test every other value of m until 9 (M = 4)
normcdfList = norm.cdf(pList) # generate 50 values for the p with std.
    normal cdf for comparison

numInt = plt.figure() # start figure...

### Do Double Numeric Integration to Estimate values for the CDF of P ###
for eachParam in mList: # do it for all M values of m
        cdfList = [] # create empty list to store estimated CDF values
        m = eachParam # shorter notation
        # estimate the constant
        const = special.gamma(m-1) * 2**(3*(1-m/2)) / (special.gamma(m/2)
            * special.gamma((m-1)/2) * special.gamma((m-1)/2) )
        for eachPoint in pList: # estimate F_P(p) for all 50 data points
            (-3 to 3)
                cdf = dblquad(lambda t, p: const * 1/abs(t) * (p/t)**(m-1)
                    * np.exp(-0.5*(p/t)**2) * (1-t**2)**((m-3)/2), -np.inf
                    , eachPoint, lambda p: 0, lambda p: 1)
                cdf = cdf[0] # extract approximation and toss error term
                cdfList.append(cdf) # append approximation to list for
                    comparison
        plotLabel = r'CDF of $P \,$ with $m=$' + str(m) # raw string for
            LaTeX
        plt.plot(pList, cdfList, label=plotLabel) # plot approx. cdf

### Finish Plot ###
plt.plot(pList, normcdfList, label=r'CDF of Std. Normal RV') # raw string
    for LaTeX
```

```
35  plt.title('CDFs Approximated Using Numeric Integration')
36  plt.legend(loc='lower right')
37  os.chdir('../Figures/') # navigate to "Figures" folder
38  plt.savefig('13.pdf') # store figure as pdf
39  plt.close(numInt)
40
41  ### Create LaTeX Table with Comparison for m = 9 ###
42  diffList = abs(cdfList - normcdfList) # list of absolute values of
        differences
43  comparisonDF = pd.DataFrame({'Input Value' : pList, 'Num. Integration' :
        cdfList, 'Normal CDF' : norm.cdf(pList), 'Absolute Difference' :
        diffList})
44  comparisonDF = comparisonDF[['Input Value', 'Num. Integration', 'Normal
        CDF', 'Absolute Difference']]
45  comparisonDF.drop(range(1,48,2), axis=0, inplace = True) # drop every
        other row (readability)
46  # export table to a .tex file that is loaded by my LaTeX document...
        reproducibility FTW
47  table = comparisonDF.to_latex(index = False)
48  writeFile = codecs.open('NumericIntegration.tex','w', 'utf-8')
49  writeFile.write(table)
50  writeFile.close()
```

14. Let $X_i \overset{iid}{\sim} \text{Unif}(0,1)$, $i = 1, \ldots, n$, let $s$ and $t$ be values such that $0 < s < t < 1$, and define $N_n(s) = \sum_{i=1}^{n} \mathbb{I}_{[0,s]}(X_i)$ such that $N_n(s)$ denotes the number of $X_i$ which are less than or equal to $s$. Furthermore, let $X = N_n(s)$, $Y = N_n(t)$, and $D = Y - X$.

**Proposition:**     $(D|X = x) \sim \text{Bin}(n - x, (t - s)/(1 - s))$

*Proof.* We're mainly interested in the distribution of $D$ (conditional on $X$) and plan to use a bivariate transformation that expresses the joint pmf of $D$ and some dummy variable $M$ as a function of the joint pmf of $X$ and $Y$. Let us thus start by obtaining the pmf of $X$. Since $X$ is just a sum of Bernoulli random variables $X_i$, it follows a binomial distribution:

$$f_X(x; n, s) = \binom{n}{x} s^x (1 - s)^{n-x} \, \mathbb{I}(0 \leq x \leq n)$$

Let us now think about the joint distribution of $X$ and $Y$. We know - by definition - that $x$ values are contained in the interval $[0, s]$ and $y$ values are within $[0, t]$. This implies that $y - x$ values are contained in the interval $(s, t]$. Finally, we conclude that $n - y$ values exceed $t$. Hence we've now identified three possibilities, which - in conjunction with the fact that all $X_i$ are iid - enables us to deduce that the joint pmf of $X$ and $Y$ must be trionomial. More specifically, we can state their joint pmf as follows:

$$f_{X,Y}(x, y; n, s, t) = \binom{n}{x, \, y - x, \, n - y} s^x (t - s)^{y-x} (1 - t)^{n-y} \, \mathbb{I}(0 \leq x \leq y \leq n)$$

Note that the equation above makes use of the trinomial coefficient, defined (in general) as follows:

$$\binom{n}{a, b, c} = \frac{n!}{a! \, b! \, c!}$$

Let us now, as alluded to previously, state the joint pmf of $D$ and $M$ (the aforementioned dummy that enables us to use a bivariate transformation) as a function of the joint pmf of $X$ and $Y$ we just derived.

$$f_{D,M}(d, m) = f_{X,Y}(x, y)|\det \mathbf{J}|$$

Note that this references the Jacobian matrix (which we are yet to state) and no longer explicitly indicates that all these functions depend on $n, s, t$, which improves readability. Recall that we've defined $D = Y - X$ and let $M$ be equal to $X$. Obviously, this is equivalent to stating $X = M$ and $Y = D + M$. We can now address the Jacobian matrix:

$$\mathbf{J} = \begin{vmatrix} \dfrac{\partial x}{\partial d} & \dfrac{\partial x}{\partial m} \\ \dfrac{\partial y}{\partial d} & \dfrac{\partial y}{\partial m} \end{vmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$|\det \mathbf{J}| = |0 \cdot 1 - 1 \cdot 1| = |-1| = 1$$

Since this evaluates to one, we can simplify our joint pmf for $D$ and $M$ by omitting the Jacobi term.

$$f_{D,M}(d, m) = f_{X,Y}(x, y)$$
$$= f_{X,Y}(m, d + m)$$

Let us now insert the expression for the joint pmf of $X$ and $Y$ we derived already.

$$f_{D,M}(d, m) = \binom{n}{m, \, d, \, n - (d+m)} s^m (t - s)^d (1 - t)^{n-(d+m)} \, \mathbb{I}(0 \leq m \leq d + m \leq n)$$

$$= \frac{n!}{m! \, d! \, (n - (d + m))!} s^m (t - s)^d (1 - t)^{n-(d+m)} \, \mathbb{I}(0 \leq m \leq d + m \leq n)$$

$$= \frac{n!}{m! \, (n - m)! \, d! \, (n - m - d)!} (n - m)! \, s^m (t - s)^d (1 - t)^{n-(d+m)} \, \mathbb{I}(0 \leq m \leq d + m \leq n)$$

$$= \binom{n}{m} \binom{n - m}{d} s^m (t - s)^d (1 - t)^{n-(d+m)} \, \mathbb{I}(0 \leq m \leq d + m \leq n)$$

Now that we have an expression for their joint pmf, recall that we're actually interested in the distribution of $D$ conditional on $X$. Since we've defined $M = X$, we can use the definition of a conditional probability and apply it to our pmfs:

$$f_{D|M}(d|m) = \frac{f_{D,M}(d, m)}{f_M(m)} = \frac{f_{D,M}(d, m)}{f_X(m)}$$

$$= \frac{\binom{n}{m} \binom{n-m}{d}}{\binom{n}{m}} \frac{\cancel{s^m}(t - s)^d (1 - t)^{n-(d+m)}}{\cancel{s^m}(1 - s)^{n-m}} \, \mathbb{I}(0 \leq m \leq d + m \leq n) \, \mathbb{I}(0 \leq m \leq n)$$

The two indicator functions can - using basic logic - be combined into one.

$$f_{D|M}(d|m) = \frac{\binom{n}{m} \binom{n-m}{d}}{\binom{n}{m}} \frac{(t - s)^d (1 - t)^{n-(d+m)}}{(1 - s)^{n-m}} \, \mathbb{I}(0 \leq d \leq n - m)$$

$$= \frac{\cancel{\binom{n}{m}}\binom{n-m}{d}}{\cancel{\binom{n}{m}}} \frac{(t-s)^d(1-t)^{n-(d+m)}}{(1-s)^{n-m}} \mathbb{I}(0 \le d \le n-m)$$

$$= \binom{n-m}{d} \frac{(t-s)^d(1-t)^{n-(d+m)}}{(1-s)^{n-m}(1-s)^d(1-s)^{-d}} \mathbb{I}(0 \le d \le n-m)$$

$$= \binom{n-m}{d} \left(\frac{t-s}{1-s}\right)^d \left(\frac{1-t}{1-s}\right)^{n-(d+m)} \mathbb{I}(0 \le d \le n-m)$$

$$= \binom{n-m}{d} \left(\frac{t-s}{1-s}\right)^d \left(1-\frac{t-s}{1-s}\right)^{(n-m)-d} \mathbb{I}(0 \le d \le n-m)$$

Or since $M = X$:

$$f_{D|X}(d|x) = \binom{n-x}{d} \left(\frac{t-s}{1-s}\right)^d \left(1-\frac{t-s}{1-s}\right)^{(n-x)-d} \mathbb{I}(0 \le d \le n-x)$$

The last expression represents the pmf of a Binomial random variable with $n-x$ trials and a probability of success of $\frac{t-s}{1-s}$. Hence we've proven that $(D|X = x)$ does indeed follow a Binomial distribution:

$$(D|X = x) \sim \text{Bin}(n-x, (t-s)/(1-s))$$

∎