Name Yaniv Bronshtein yb262

M LuValle, Fall 2021 581 Midterm 1

1) Let X1 and X2 be two independent N(0,1), (normal mean 0, variance 1) random variables. Let $Y = X_1^2 + X_2^2$.

a. What are the two names for the distribution of Y?

$Y = \sum_{i=1}^{n} X_i^2$ Therefore $Y \sim \chi^2(2)$ or a Chi-Squared Distribution with 2 degrees of freedom. This also a special case of the Gamma Distribution

b. Calculate (integrate out) the mean and variance of Y: Show your work.

$f(x) = \frac{1}{\Gamma(p/2)\, 2^{p/2}} x^{p/2-1} e^{-x/2}, \quad x>0$

Note: $\Gamma(n) = (n-1)!$
$\Gamma(1) = 0! = 1$

Plug in $p=2$

$f(x) = \frac{1}{\Gamma(1) * 2} x^0 e^{-x/2}$

$f(x) = \frac{1}{2} e^{-x/2}$

$E[x] = \frac{1}{2}\int_0^\infty x * e^{-x/2} dx =$

By Parts:

Recall $\int u\,dv = uv - \int v\,du$

Let $u = \frac{x}{2} \Rightarrow \frac{du}{dx} = \frac{1}{2}$ $\quad 2du = dx$

$du = \frac{1}{2} dx$

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

$v = \int e^{-\frac{x}{2}} dx = -2e^{-x/2}$

Thus:

$E[Y] = \frac{x}{2} * (-2e^{-x/2}) - \int -2e^{-x/2} \frac{1}{2} dx =$

$= -xe^{-x/2} - 2e^{-x/2} = -(x+2)e^{-x/2}$

Let us apply the proper limits

$$\left[-(x+2)e^{-x/2}\right]_0^{\infty} = \left[(x+2)e^{-x/2}\right]_{\infty}^{0} = 2-0=2$$

Thus, $\boxed{E[Y]=2}$

$$Var[Y] = E[Y^2]-(E[Y])^2$$

Find $E[Y^2]$

① $E[Y^2] = \int_0^{\infty} \frac{1}{2}x^2 e^{-x/2}\,dx$

Use same $u=x^2 \Rightarrow du=2x\,dx$

$dv=e^{-x/2}dx \Rightarrow v=\int e^{-x/2}dx = -2e^{-\frac{x}{2}}$

$$E[Y^2] = -e^{-x/2}x^2+2\int e^{-x/2}x\,dx$$

Again apply integration by parts

This time, $u=x \Rightarrow du=dx \quad v=-2e^{-x/2}$

So we get:

$$-4e^{-x/2}x-e^{-x/2}x^2+4\int e^{-x/2}dx=$$

$$=-e^{-x/2}x^2-4e^{-x/2}x-8e^{-x/2}+C$$

Evaluate at bounds $[0,\infty)$

$$=\left[e^{-x/2}x^2+4e^{-x/2}x+8e^{-x/2}\right]_{\infty}^{0}$$ change bounds

$= 8 - 0 = 8$

② Now find Var[Y]

$$\text{Var}[Y] = E[Y^2] - (E[Y])^2$$
$$= 8 - (2)^2 = \boxed{4}$$

2) Let X_i be iid random variables with distribution $X_i = \begin{cases} -1 \; with \; probability \; P \\ 1 \; with \; probability \; 1-p \end{cases}$. Let $Y_n = \sum_{i=1}^{n} X_i$.

a) Calculate the mean and variance of Y_n

$E[X_i] = -1 * P + 1(1-P) = 1-2p$

$E[X_i^2] = (-1)^2 * p + (1)^2 * (1-p)$

$= 1$

$Var[X_i] = E[X_i^2] - (E[X_i])^2$

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

$= 1 - (1 - 4p - 4p)$

$= 4p(1-p)$

Thus,

$\boxed{E[Y_n] = n * (1-2p) \text{ and } Var[Y_n] = n * 4p(1-p)}$

b) and the covariance of $Y_n Y_m, m < n$.

Recall:

$$Cov\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{} \sum_{j=1}^{} a_i b_j Cov(X_i, Y_j)$$

In our case let all $a_i = 1$ and all $b_j = 1$

Then,

$$Cov\left(\sum_{i=1}^{m} X_i, \sum_{j=1}^{n} X_j\right) = \sum_{j=1}^{n} \sum_{i=1}^{m} cov(X_i, X_j) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} E[X_i X_j] - E[X_i] E[X_j]$$

Since we are dealing with I.I.D random variables we have two cases:

Case 1:

$i=j$

Then we have $Cov(X_i, X_i) = Var(X_i)$

Case 2:

$i \neq j$

Then we have two independent vectors

$E[X_i X_j] = E[X_i] * E[X_j]$

Thus $Cov(X_i, X_j) = E[X_i]E[X_j] - E[X_i]E[X_j]$

We can use a matrix to visualize.

$$
\begin{array}{c}
\overbrace{\hphantom{Var(X_1) \quad 0 \quad \cdots \quad 0 \quad Var(X_m)}}^{m} \\
\left[\begin{array}{ccccc}
Var(X_1) & 0 & \cdots & \cdots & 0 \\
0 & Var(X_2) & 0 & & \vdots \\
\vdots & 0 & Var(X_3) & & \vdots \\
\vdots & & 0 & \ddots & 0 \\
0 & \cdots & \cdots & 0 & Var(X_m) \\
0 & \cdots & \cdots & \cdots & 0 \\
\vdots & & & & \vdots \\
0 & \cdots & \cdots & \cdots & 0
\end{array}\right]
\begin{array}{l}
\left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\end{array}}\right\} m \\
\left.\vphantom{\begin{array}{c}1\\1\\1\end{array}}\right\} n-m
\end{array}
\end{array}
$$

$Var(X_i) = 4p(1-p)$

$$\boxed{\begin{aligned} &\text{So } Cov(Y_n, Y_m) = \\ &= m * 4p(1-p) \end{aligned}}$$

3) Using Y_n from problem 2, show that

a. ~~Y_n/n converges in probability to $E(Y_n)$~~ $\frac{Y_n}{n} \rightarrow \frac{E(Y_n)}{n}$

See solutions on next Page

b. $\frac{Y_n - E(Y_n)}{\sqrt{n}}$ is asymptotically normal with what variance?

c. Y_m and $Y_n - Y_m$, m<n are independent

3a).

From the previous question:

$E[Y_n] = n(1-2p)$

$Var[Y_n] = 4n(1-p)p$

The Weak Law Of Large Numbers (WLLN) states that if $X_1, \ldots, X_n$ are I.I.D then $\overline{X_n} \xrightarrow{P} \mu$

Thus,

$$\frac{X_1 + X_2 + X_3 + \ldots + X_n}{n} \xrightarrow{P} \mu$$

Replacing $X_1 + \ldots + X_n$ with Y_n,

$$\frac{Y_n}{n} \xrightarrow{P} \mu$$

where $\mu = \frac{n(1-2p)}{n} = 1-2p$

b). The Central Limit Theorem states that if $\overline{X_n} = \frac{Y_n}{n}$ then

$$Z_n = \frac{\frac{Y_n}{n} - \mu}{\sqrt{Var(\frac{Y_n}{n})}} = \frac{\frac{Y_n}{n} - \mu}{\sqrt{\frac{1}{n^2} * n * 4p(1-p)}} = \frac{\frac{Y_n}{n} - (1-2p)}{\frac{1}{\sqrt{n}} * \sqrt{4p(1-p)}} = \frac{\frac{Y_n - n(1-2p)}{n}}{\frac{1}{\sqrt{n}}\sqrt{4p(1-p)}}$$

$$= \frac{Y_n - n(1-2p)}{\sqrt{4np(1-p)}} =$$

$$= \frac{Y_n - n(1-2p)}{\sqrt{n} * \sqrt{4p(1-p)}}$$

Observe that Z_n converges to $Z \sim N(0,1)$.

Multiplying Z_n by $\sqrt{4p(1-p)}$ generates $\frac{Y_n - E[Y_n]}{\sqrt{n}}$

Thus, the standard deviation of 1 is multiplied by that quantity.

Taking the square, we conclude that our distribution is asymptotically normal with variance $4p(1-p)$.

c). $Y_m = \sum_{i=1}^{m} X_i$

Let's write Y_n using m.

$$Y_n = \sum_{i=1}^{n} X_i = \sum_{i=1}^{m} X_i + \sum_{i=m+1}^{n} X_i$$

$$Y_n - Y_m = \sum_{i=m+1}^{n} X_i$$

Recall the Variance

1 $\text{Var}[Y_m] = 4mp(1-p)$

$$\text{Var}[Y_n - Y_m] = \text{Var}[X_{m+1} + X_{m+2} + X_{m+3} + \dots + X_n]$$
$$= (n-m)4p(1-p)$$

2 $\text{Var}[Y_m] + \text{Var}(Y_n - Y_m) = 4mp(1-p) + 4(n-m)p(1-p)$

$$= 4mp(1-p) + 4np(1-p) - 4mp(1-p)$$

cancel out terms

$$= 4np(1-p)$$

③ $Var[Y_m + Y_{n-m}] = Var\left[\sum_{i=1}^{m} X_i + \sum_{j=m+1}^{n} X_j\right] = Var\left[\sum_{i=1}^{n} X_i\right]$

$= 4np(1-p)$

Thus $Var[Y_m] + Var[Y_n - Y_m] = Var[Y_m + Y_{n-m}]$
and we have proven independence ∎

Poisson

4) You have been put in charge of defending the earth from an asteroid strikes. Asteroid strikes are rare events where the frequency of impact scales with the size according to a Poisson distribution. For calibration, the meteor over Chelyabinsk in 2013 detenoted with an airburst equivalent to 500 Kilotons of TNT (30 times the power of the bomb over Hiroshima) and injured 500 people, it was 20 meters in diameter, and will occur with a frequency of about 1 in 60 years (but remember, most of the earths surface is water). A Hiroshima size event occurs about every year. A 70 meter asteroid will result in an airburst with an energy of 1.9 megatons of TNT, and will occur about once every 1900 years. Assume you are given a threshold of 70 meters as what you have to defend against, and you have to defend earths cities which cover 3% of earths surface. Assume 12% to account for strikes which could affect cities. Your job starts by getting a budget.

a. What is the rate of impact for 12% of the earths surface.

$$\text{Rate of Impact} = \frac{0.12}{1900} = 0.000063158$$

b. It is estimate 29% of the earth is habitable, what is the impact rate rate for the habital earths surface.

If 29% of Earth is habitable:

$$0.29 * \frac{1}{1900} = 0.00015263$$

c. What is the probability of 1 or more impact events you will be responsible for during your life time (assume 100 years optimistically, 12% of surface).

See Next Page

d. How does this translate to expected yearly loss in life and in money during your lifetime if an average city has 1 million people and Annual GDP for an average city is 64 billion dollars.

See Next Page(s)

c). The Probability of One or more Impacts is determined using Poisson pdf

$P(\#\text{Impacts} \geq 1) = 1 - P(\#\text{Impacts} = 0) =$

$$= 1 - \frac{e^{-0.0063158} * (0.0063158)^0}{0!} = 1 - 0.993704 = \boxed{0.0062959}$$

$$\lambda = \frac{0.12*100}{1900} = 0.0063158$$

d). $\text{Expected loss of Life} = \frac{0.0063158}{100} * 10^6 \approx \boxed{63.158 \ \frac{\text{lives}}{\text{yr}}}$

Expect Total Loss of Money =

$$= 64*10^9 \text{ dollars} * 0.0063158 = \boxed{\$404{,}211{,}200}$$

Hint: Figure out Limiting Distribution

5) Let Xn be binomial n with probability p=1/n, as n-> infinity.

a. What does the distribution of Xn converge to?

Binomial(n,p) defined as $f(x)=\binom{n}{x}p^x(1-p)^{n-x}$ for $x\in\{0,1,\dots,n\}$

$f(x)=\binom{n}{x}\left(\frac{1}{n}\right)^x\left(\frac{n-1}{n}\right)^{n-x}$

Binomial$(n,\frac{\lambda}{n})$ converges to Poisson(λ) defined by $f(x)=\frac{e^{-\lambda}\lambda^x}{x!}$ for $x\in\{0,1,\dots\}$

so Binomial$(n,\frac{1}{n})$ converges to $\frac{e^{-1}1^x}{x!}$ Poisson$(\lambda=1)$

b. So if we add a number of such Xn together, what is the distribution of that random variable?

If we have $X_1\sim$ Poisson(λ_1) and $X_2\sim$ Poisson(λ_2) and $X_3=X_1+X_2$ then $X_3\sim$ Poisson$(\lambda_1+\lambda_2)$.

By that logic let m be some integer >2

$m*X_n\sim$ Poisson$(m\lambda)$

c. So if you have a coin with probability π/1000 of coming up heads and you make 1000 tosses use the result above to estimate the probability of 4 or more heads.

Let $P(H)=\frac{\pi}{1000}$

Then $\lambda=np=1000*\frac{\pi}{1000}=\pi$

$P(X\geq 4)=1-P(X<4)=1-[P(X=0)+P(X=1)+P(X=2)+P(X=3)]$

$=1-\left[\frac{e^{-\pi}\lambda^0}{0!}+\frac{e^{-\pi}\lambda^1}{1!}+\frac{e^{-\pi}\lambda^2}{2!}+\frac{e^{-\pi}\lambda^3}{3!}\right]=1-\left[e^{-\pi}\left(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{6}\right)\right]$

$=1-\left[e^{-\pi}\left(1+\pi+\frac{\pi^2}{2}+\frac{\pi^3}{6}\right)\right]\approx \boxed{0.384456}$

6) Consider the random variable $Z_n = cos\left(2 * \pi * {}^{Y_n}/_{n}\right)$ where Yn is binomial(n,p)

a. What is its asymptotic variance for arbitrary p as n increases

See Next Page

b. What happens at p=1/2

See Next Page

c. Write an r program that shows you graphical the (simulated) standard deviation of Z_n vs 1/sqrt(n) as n increases for different values of p including 0.5

Write the code here:

a). Let $Z_n = \cos(2\pi * \frac{Y_n}{n})$

where $Y_n \sim \text{Binomial}(n,p)$

As $n \to \infty$, $Y_n \sim N(\mu = np, \sigma^2 = np(1-p))$

Let us apply the Delta Method:

$$g(X_n) \approx N(g(\mu), \frac{g'(\mu)^2 \sigma^2}{n})$$

Let $g(\mu) = \cos(2\pi\mu) \Rightarrow g(\mu = np) = \cos(2\pi np)$

$g'(\mu) = -2\pi \sin(2\pi\mu) \Rightarrow g'(\mu = np) = -2\pi * \sin(2\pi * np)$

$$g(X_n) \approx N(\cos(2\pi np), \frac{(-2\pi * \sin(2\pi np))^2 * np(1-p)}{n})$$

$$g(X_n) \approx N(\cos(2\pi np), 4\pi^2 * \sin^2(2\pi np) * p(1-p))$$

Therefore, $\boxed{Var = 4\pi^2 \sin^2(2\pi np) * p(1-p)}$

b). At $p = \frac{1}{2}$

$$Var = 4\pi^2 \sin^2(2\pi n * \tfrac{1}{2}) * \tfrac{1}{2}(1 - \tfrac{1}{2})$$

Since $\sin(\pi) = 0$, the whole expression $= 0$

$\boxed{Var \text{ at } p = \frac{1}{2} = 0}$