

1 PHY509 – Problem Set 2

Problem 1

Let X, Y be independent random variables. The Conditional Expectation Formula is:

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y))$$

To be clearer, we can explicitly denote which variable the expectation is with respect to as follows:

$$\mathbb{E}_X(X) = \mathbb{E}_Y(\mathbb{E}_X(X|Y))$$

- (a) Prove this for the case of discrete X, Y .
 - (b) Suppose you are playing a computer game in which you are stuck in a room with 3 exits that are indistinguishable, which means you have equal probability of choosing exit 1, 2 or 3. If you leave via exit 1, you get free after traversing 3 other rooms. If you leave via exit 2, you end up back at this same starting room after traversing 5 rooms. If you leave via exit 3, you end up back at this same starting room after traversing 7 rooms. Sadly, if you get back to the room, you are so disoriented that you can't tell which exit you chose last time, so you randomly pick again. How many rooms do you visit on average before getting free? (Bonus: how annoying is the game designer?)
-

- (a) Let's prove this by showing that the right hand side equals the left hand side and also using the definition of expectation value for discrete random variables.

$$E(X) = \sum_i x_i P(X = x_i)$$

We want to expand out the inner expectation value first:

$$E_X(X|Y = y) = \sum_i x_i P(X = x_i | Y = y)$$

Plugging that back into the outer expectation value and also using the definition of expectation value for discrete random variables again:

$$E_Y(E_X(X|Y)) = \sum_j \left(\sum_i x_i P(X = x_i | Y = y_j) \right) P(Y = y_j)$$

From conditional probability we have:

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \quad \text{for } P(F) > 0$$

$$P(E|F)P(F) = P(E \cap F) = P(E, F)$$

Such that we can write our expectation expression in terms of the joint probability:

$$E_Y(E_X(X|Y)) = \sum_j \sum_i x_i P(X = x_i, Y = y_j)$$

For independent events (like for the case of X and Y here) we have $P(A \cap B) = P(A)P(B)$ so that:

$$E_Y(E_X(X|Y)) = \sum_j \sum_i x_i P(X = x_i) P(Y = y_j)$$

We can move over the terms to their respective sums:

$$E_Y(E_X(X|Y)) = \sum_i x_i P(X = x_i) \sum_j P(Y = y_j)$$

From the laws of probability we have the law of total probability which states that the sum of the probabilities of all possible outcomes of a random variable is equal to 1, which means that $\sum_j P(Y = y_j) = 1$.

$$E_Y(E_X(X|Y)) = \sum_i x_i P(X = x_i) \cdot 1 = \sum_i x_i P(X = x_i) = E(X)$$

(b) Let's let X be the number of rooms visited before getting free and Y be the exit chosen.

Just to get the problem setup let's remind ourselves of the exit options from the starting room:

- Exit 1: 3 rooms to freedom
- Exit 2: 5 rooms then back to start
- Exit 3: 7 rooms then back to start

The question is asking for the expected number of rooms visited before getting free, which is $E(X)$. So using the problem 1 setup let's use the conditional expectation formula we just proved and see if it can help us break down the problem:

$$E_X(X) = E_Y(E_X(X|Y))$$

Putting this into the definition of expectation value for discrete random variables:

$$E_X(X) = \sum_j \sum_i x_i P(X = x_i | Y = y_j) P(Y = y_j)$$

Since there are 3 exits that are indistinguishable we have $P(Y = y_j) = \frac{1}{3}$ for $j = 1, 2, 3$.

$$E_X(X) = \frac{1}{3} \sum_j \sum_i x_i P(X = x_i | Y = y_j)$$

Putting the expectation value notation back in here will make it easier.

$$E_X(X) = \frac{1}{3} \sum_j E_X(X|Y = y_j)$$

Now we need to figure out $E_X(X|Y = y_j)$ for each exit choice j .

- For exit 1 we have $E_X(X|Y = 1) = 3$ since we just go through 3 rooms and get free.
- For exit 2 we have $E_X(X|Y = 2) = 5 + E_X(X)$ since we go through 5 rooms and then end up back at the starting room.

- For exit 3 we have $E_X(X|Y = 3) = 7 + E_X(X)$ using the same reasoning as exit 2.

Plugging this all in to our expression we have:

$$E_X(X) = \frac{1}{3} (3 + (5 + E_X(X)) + (7 + E_X(X)))$$

$$E_X(X) = \frac{1}{3} (15 + 2E_X(X))$$

Rearranging we have:

$$3E_X(X) - 2E_X(X) = 15$$

$$E_X(X) = 15$$

So on average we visit 15 rooms before getting free. The game designer is very annoying!

Problem 2

Calculate the characteristic function of a Gaussian of variance σ^2 and show it is a Gaussian of variance $\frac{1}{\sigma^2}$.

The characteristic function is defined as:

$$\phi(t) = \mathbb{E}(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} p(x) dx$$

We can the plug in the Gaussian pdf with mean μ and variance σ^2 :

$$p_{\text{Gaussian}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

I used a LLM to generate MATLAB code to solve the integral symbolically:

```
% Define symbolic variables
syms x t mu sigma real
assume(sigma > 0)

% Define the integrand
integrand = exp(1i*t*x) * (1/sqrt(2*pi*sigma^2)) * exp(-(x-mu)^2/(2*sigma^2));

% Compute the integral from -inf to +inf
result = int(integrand, x, -inf, inf);

% Simplify the result
result_simplified = simplify(result);
```

```
% Display the result
disp('Result:')
disp(result_simplified)

% Convert to LaTeX
latex_result = latex(result_simplified);
disp(' ')
disp('LaTeX code:')
disp(latex_result)

% Pretty print in MATLAB
disp(' ')
disp('Pretty form:')
pretty(result_simplified)
```

This gives the result:

$$\phi(t) = e^{-\frac{\sigma^2 t^2}{2} + \mu t \text{li}}$$

This is a Gaussian (in t) with variance $\frac{1}{\sigma^2}$ and mean μ and an additional complex exponential factor. And if the mean is $\mu = 0$ then the imaginary term disappears and we get just $\phi(t) = e^{-\frac{\sigma^2 t^2}{2}}$

Problem 3

- (a) Show that, if the characteristic functions for two independent random variables X, Y are $\phi_X(t)$ and $\phi_Y(t)$, then the characteristic function for the random variable made from the sum, $Z = X + Y$, is just the product $\phi_Z(t) = \phi_X(t)\phi_Y(t)$.
- (b) If X_1, X_2 are two independent Gaussian random variables with means μ_1, μ_2 and variances σ_1^2, σ_2^2 , prove that the random variable $Y = a_1 X_1 + a_2 X_2$ ($a_i \in \mathbb{R}$) is Gaussian with mean $\mu = a_1 \mu_1 + a_2 \mu_2$ and variance $\sigma^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2$.

-
- (a) We start with the definition of the characteristic function for Z :

$$\phi_Z(t) = \mathbb{E}(e^{itZ}) = \mathbb{E}(e^{it(X+Y)})$$

This can be rewritten as:

$$\phi_Z(t) = \mathbb{E}(e^{itX} e^{itY})$$

Since X and Y are independent random variables, so we can also say that e^{itX} and e^{itY} are independent. Thus, from the independence condition, we can factor the expectation:

$$\phi_Z(t) = \mathbb{E}(e^{itX})\mathbb{E}(e^{itY}) = \phi_X(t)\phi_Y(t)$$

Thus, we have shown that:

$$\phi_Z(t) = \phi_X(t)\phi_Y(t)$$

- (b) From Problem 2, we know that the characteristic function of a Gaussian random variable X_i with mean μ_i and variance σ_i^2 is given by:

$$\phi_{X_i}(t) = e^{-\frac{\sigma_i^2 t^2}{2} + i\mu_i t}$$

We can then consider the random variable $Y = a_1 X_1 + a_2 X_2$. The characteristic function of Y is:

$$\phi_Y(t) = \mathbb{E}(e^{itY}) = \mathbb{E}(e^{it(a_1 X_1 + a_2 X_2)}) = \mathbb{E}(e^{ia_1 t X_1} e^{ia_2 t X_2})$$

Using the independence of X_1 and X_2 , we can factor this expectation:

$$\phi_Y(t) = \mathbb{E}(e^{ia_1 t X_1}) \mathbb{E}(e^{ia_2 t X_2}) = \phi_{X_1}(a_1 t) \phi_{X_2}(a_2 t)$$

Substituting the characteristic functions for X_1 and X_2

$$\phi_Y(t) = e^{-\frac{\sigma_1^2 (a_1 t)^2}{2} + i\mu_1 (a_1 t)} e^{-\frac{\sigma_2^2 (a_2 t)^2}{2} + i\mu_2 (a_2 t)}$$

Combining the exponents, we get:

$$\phi_Y(t) = e^{-\frac{(a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2) t^2}{2} + i(a_1 \mu_1 + a_2 \mu_2) t}$$

This is the characteristic function of a Gaussian random variable with mean $\mu = a_1 \mu_1 + a_2 \mu_2$ and variance $\sigma^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2$, which also means that the random variable $Y = a_1 X_1 + a_2 X_2$ ($a_i \in \mathbb{R}$) is

$$Y \sim \mathcal{N}(a_1 \mu_1 + a_2 \mu_2, a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2)$$

Problem 4

- (a) Show that the characteristic function for a Poisson is

$$\phi(t) = \sum_{r=0}^{\infty} e^{itr} P(r; \mu) = \exp(\mu(e^{it} - 1))$$

and check that you get the right mean and variance from the characteristic function.

- (b) Use this to show that the sum of two Poisson random variables with means μ_X, μ_Y is a Poisson with mean $\mu_X + \mu_Y$.
- (c) Prove this the hard way, by computing this directly from the formulas for the Poissons. Note how much easier it was to prove using the characteristic functions!

- (a) Starting with the definition of the characteristic function

$$\phi_X(t) = E[e^{itX}] = \begin{cases} \int_{-\infty}^{\infty} e^{itx} f_X(x) dx & \text{for continuous } X \\ \sum_k e^{itx_k} p_X(x_k) & \text{for discrete } X \end{cases}$$

Recall the Poisson distribution definitions:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$E[X] = \lambda, \quad \text{Var}(X) = \lambda$$

The discrete Poisson random variable for this question has the PMF of:

$$P(r; \mu) = \frac{\mu^r e^{-\mu}}{r!}$$

So our characteristic function becomes

$$\phi(t) = \sum_{r=0}^{\infty} e^{itr} P(r; \mu)$$

We have:

$$\phi(t) = \sum_{r=0}^{\infty} e^{itr} \frac{\mu^r e^{-\mu}}{r!} = e^{-\mu} \sum_{r=0}^{\infty} \frac{(\mu e^{it})^r}{r!}$$

recall the Taylor series expansion for e^x :

$$e^x = \sum_{r=0}^{\infty} \frac{x^r}{r!}$$

We can rewrite the characteristic function as:

$$\boxed{\phi(t) = e^{-\mu} e^{\mu e^{it}} = e^{\mu(e^{it}-1)}}$$

Recall that the one of the properties of characteristic functions is that the n th moment of the distribution can be found by taking the n th derivative of the characteristic function and evaluating it at $t=0$:

$$E[X^n] = \frac{1}{i^n} \phi_X^{(n)}(0)$$

To find the mean and variance, we then just compute the first and second derivatives of $\phi(t)$ at $t = 0$:

I used a LLM to generate MATLAB code to compute the derivatives symbolically:

```
% Define symbolic variables
syms t mu real
assume(mu > 0)

% Define the characteristic function for Poisson
phi = exp(mu*(exp(1i*t) - 1));

% Compute first derivative
phi_1 = diff(phi, t);

% Evaluate first derivative at t=0
phi_1_at_0 = subs(phi_1, t, 0);
```

```
% Compute second derivative
phi_2 = diff(phi, t, 2);
```

That MATLAB code gives me the following equations:

$$\phi(t) = e^{\mu(e^{it} - 1)}$$

$$\phi'(t) = \mu e^{it} e^{\mu(e^{it} - 1)} i$$

$$\phi'(0) = \mu i$$

$$\phi''(t) = -\mu^2 e^{it} e^{\mu(e^{it} - 1)} - \mu e^{it} e^{\mu(e^{it} - 1)}$$

$$\phi''(0) = -\mu^2 - \mu$$

The mean is defined as $E[X]$ which is the first moment, and variance is defined as $V(X) = E[X^2] - (E[X])^2$.

So we have:

$$E[X] = \frac{1}{i} \phi'(0) = \frac{1}{i} (\mu i) = \mu$$

And for the variance:

$$E[X^2] = \frac{1}{i^2} \phi''(0) = -\phi''(0) = -(-\mu^2 - \mu) = \mu^2 + \mu$$

Thus:

$$V(X) = E[X^2] - (E[X])^2 = (\mu^2 + \mu) - (\mu)^2 = \mu$$

So we have confirmed that the mean and variance of the Poisson distribution are both equal to μ .

- (b) From question 3 (a) we found that if the characteristic functions for two independent random variables X , Y are $\phi_X(t)$ and $\phi_Y(t)$, then the characteristic function for the random variable made from the sum, $Z = X + Y$, is just the product:

$$\phi_Z(t) = \phi_X(t) \phi_Y(t)$$

Recall from part (a) of this question that the characteristic function for a Poisson random variable with mean μ is:

$$\phi(t) = e^{\mu(e^{it} - 1)}$$

Such that the characteristic function for the sum of two independent Poisson random variables with means μ_X and μ_Y is:

$$\phi_Z(t) = \phi_X(t)\phi_Y(t) = e^{\mu_X(e^{it}-1)}e^{\mu_Y(e^{it}-1)} = e^{(\mu_X+\mu_Y)(e^{it}-1)}$$

This being the characteristic function of a Poisson random variable with mean $\mu_X + \mu_Y$.

$$Z \sim P(r; \mu_X + \mu_Y)$$

- (c) To now prove that the sum of two Poisson random variables with means μ_X, μ_Y is a Poisson with mean $\mu_X + \mu_Y$. the hard way, we will be computing this directly from the formulas for the Poissons:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$E[X] = \lambda, \quad \text{Var}(X) = \lambda$$

Here in this question we have $P_X(r; \mu_X) = \frac{\mu_X^r e^{-\mu_X}}{r!}$ and $P_Y(r; \mu_Y) = \frac{\mu_Y^r e^{-\mu_Y}}{r!}$. We can then write the PMF for the sum of two independent Poisson random variables X and Y as:

To find the PMF of the sum of two independent Poisson random variables X and Y , we must use the convolution formula. For independent random variables X and Y , the distribution of their sum $Z = X + Y$ is:

$$P(Z = k) = \begin{cases} \sum_{r=0}^k P(X = r) \cdot P(Y = k - r) & \text{discrete} \\ \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx & \text{continuous} \\ \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy & \text{also continuous} \end{cases}$$

So here we have:

$$P_Z(r) = \sum_{j=0}^r P_X(j; \mu_X) P_Y(r - j; \mu_Y) = \sum_{j=0}^r \frac{\mu_X^j e^{-\mu_X}}{j!} \cdot \frac{\mu_Y^{r-j} e^{-\mu_Y}}{(r-j)!}$$

$$P_Z(r) = e^{-(\mu_X + \mu_Y)} \sum_{j=0}^r \frac{\mu_X^j \mu_Y^{r-j}}{j! (r-j)!}$$

The binomial theorem tells us that:

$$\sum_{j=0}^n \binom{n}{j} a^j b^{n-j} = (a + b)^n \quad \text{where} \quad \binom{n}{j} = \frac{n!}{j!(n-j)!}$$

So we can rewrite our PMF expression as:

$$P_Z(r) = \frac{(\mu_X + \mu_Y)^r e^{-(\mu_X + \mu_Y)}}{r!}$$

this is the form of a Poisson PMF, which also shows that the sum of two Poisson random variables with means μ_X, μ_Y is a Poisson with mean $\mu_X + \mu_Y$.

Problem 5

Let $\{x_i\}$, $i = 1, \dots, n$ be a sample of size n drawn from a probability density function $p(x)$ with mean μ and variance σ^2 . Define an estimator by the sample mean

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

- (a) Show $\hat{\mu}$ is an unbiased estimator for μ .
 - (b) Show that the variance of the sample, $\hat{s}^2 = \frac{1}{n} \sum_i (x_i - \hat{\mu})^2$ is a biased estimator for the true pdf (population) variance σ^2 .
 - (c) Show that $\hat{s}^2 = \frac{1}{n-1} \sum_i (x_i - \hat{\mu})^2$ is an unbiased estimator for the true pdf variance σ^2 . (The way to remember this is that estimating the mean is one combination of the n data points, so you effectively only have $n - 1$ left to estimate the variance.)
-

- (a) To show that an estimator is unbiased, we need to show that its expectation is equal to the true value (i.e. $\mathbb{E}[\hat{\theta}] = \theta$). So we can start with the definition of the sample mean estimator:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

and show that $\hat{\mu}$ is an unbiased estimator for μ by computing the expectation value of $\hat{\mu}$ and showing that it equals μ :

$$\mathbb{E}[\hat{\mu}] = \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n x_i \right]$$

Recall that one of the properties of expectation value is linearity of expectation, which states that for any random variable X and Y , and a constant a :

$$\mathbb{E}[aX + Y] = a\mathbb{E}[X] + \mathbb{E}[Y]$$

Using this property, we can move the expectation inside the sum:

$$\mathbb{E}[\hat{\mu}] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[x_i]$$

Since each x_i is drawn from the same distribution with mean μ , we have $\mathbb{E}[x_i] = \mu$ for all i . Therefore:

$$\mathbb{E}[\hat{\mu}] = \frac{1}{n} \sum_{i=1}^n \mu = \frac{1}{n} (n\mu) = \mu$$

Thus, we have shown that:

$$\boxed{\mathbb{E}[\hat{\mu}] = \mu}$$

- (b) Following the proof in [proofwiki](#) we can show that the sample variance estimator $\hat{s}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2$ is a biased estimator for the true variance σ^2 (i.e. not equal to).

$$\begin{aligned}
\mathbb{E}[\hat{s}^2] &= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2 \right] \\
&= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n ((x_i - \mu) - (\hat{\mu} - \mu))^2 \right] \\
&= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \left((x_i - \mu)^2 - 2(x_i - \mu)(\hat{\mu} - \mu) + (\hat{\mu} - \mu)^2 \right) \right] \\
&= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 - \frac{2}{n} (\hat{\mu} - \mu) \sum_{i=1}^n (x_i - \mu) + \frac{1}{n} (\hat{\mu} - \mu)^2 \sum_{i=1}^n 1 \right]
\end{aligned}$$

We can simplify the summation term $\sum_{i=1}^n (x_i - \mu)$ as $n(\hat{\mu} - \mu)$. Substituting this in gives:

$$\begin{aligned}
&= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 - \frac{2}{n} (\hat{\mu} - \mu) \cdot n(\hat{\mu} - \mu) + (\hat{\mu} - \mu)^2 \right] \\
&= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 - 2(\hat{\mu} - \mu)^2 + (\hat{\mu} - \mu)^2 \right] \\
&= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 - (\hat{\mu} - \mu)^2 \right]
\end{aligned}$$

$$\text{Var}(\hat{\mu}) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(x_i) = \frac{1}{n^2} (n\sigma^2) = \frac{\sigma^2}{n}.$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[(x_i - \mu)^2 \right] - \mathbb{E} \left[(\hat{\mu} - \mu)^2 \right] \quad \text{(Linearity of Expectation)}$$

$$= \frac{1}{n} \sum_{i=1}^n \text{Var}(x_i) - \text{Var}(\hat{\mu}) \quad \text{(Definition of Variance)}$$

$$= \frac{1}{n} \sum_{i=1}^n \sigma^2 - \frac{\sigma^2}{n} \quad \text{(Variance of Sample Mean is } \sigma^2/n \text{)}$$

$$\begin{aligned}
&= \frac{1}{n} (n\sigma^2) - \frac{\sigma^2}{n} \\
&= \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2
\end{aligned}$$

Since $\mathbb{E}[\hat{s}^2] = \frac{n-1}{n} \sigma^2 \neq \sigma^2$, the estimator \hat{s}^2 is biased.

- (c) Since for part b for the old estimator $\hat{s}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2$ we found that $\mathbb{E}[\hat{s}^2] = \frac{n-1}{n} \sigma^2$, we can use this to say that we will need to multiply the estimator by $\frac{n}{n-1}$ to make it unbiased (i.e. to force $\mathbb{E}[\hat{s}^2] = \sigma^2$). We can then define our new estimator:

$$\boxed{\hat{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2}$$

We can also more rigorously prove that this new estimator is unbiased by computing its expectation value:

$$\mathbb{E}[\hat{s}^2] = \mathbb{E}\left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2\right]$$

And then the same steps as in part (b) but now factored by $\frac{n}{n-1}$.

Problem 6

Suppose you have a beam of particles with an average rate R per second passing through your detector.

- What is the probability of getting n particles passing through in T seconds?
- Suppose your detector has an efficiency of ϵ to successfully detect a particle. Derive the probability of getting r counts from the detector in T seconds.
- Suppose $RT = 100$ and $\epsilon = 0.10$. What is the probability of having 110 particles in this time? What is the probability of observing 11 in your detector? What is the probability of having any number between 90 and 110 particles (inclusive) in this time? What is the probability of observing between 9 and 11 in your detector?

- The number of particles passing through the detector in time T follows a Poisson distribution with mean $\mu = RT$. The probability of getting n particles is given by:

$$P(n; \mu) = \frac{\mu^n e^{-\mu}}{n!}$$

The probability of getting n particles passing through in T seconds is then:

$$P(n; RT) = \frac{(RT)^n e^{-RT}}{n!}$$

- We can see that with a detector efficiency of ϵ , the mean number of detected particles in time T is just $\mu' = \epsilon RT$. The probability of getting r counts from the detector in T seconds is then given by the Poisson distribution with this modified mean:

$$P(r; \mu') = \frac{(\mu')^r e^{-\mu'}}{r!}$$

The probability of getting r counts from the detector in T seconds is then:

$$P(r; \epsilon RT) = \frac{(\epsilon RT)^r e^{-\epsilon RT}}{r!}$$

To derive this more explicitly, we can see that the detection process is a two-step process:

- First, n particles arrive at the detector in time T , which follows a Poisson distribution with mean RT .

$$P(N = n) = \frac{(RT)^n e^{-RT}}{n!}$$

- Then, each of these n particles is detected independently with probability ϵ , which follows a Binomial distribution.

$$P(R = r | N = n) = \binom{n}{r} \epsilon^r (1 - \epsilon)^{n-r}$$

We then can use the law of total probability ($P(A) = \sum_n P(A|B_n)P(B_n)$) to find the overall probability of detecting r particles:

$$P(R = r) = \sum_{n=r}^{\infty} P(R = r | N = n) P(N = n)$$

Substituting in the expressions we found earlier:

$$P(r) = \sum_{n=r}^{\infty} \binom{n}{r} \epsilon^r (1 - \epsilon)^{n-r} \frac{(RT)^n e^{-RT}}{n!}$$

where $\binom{n}{r} = \frac{n!}{r!(n-r)!}$. Simplifying this expression, we find:

$$P(r) = \sum_{n=r}^{\infty} \frac{\epsilon^r (1 - \epsilon)^{n-r}}{r!(n-r)!} (RT)^n e^{-RT}$$

Taking out the terms that do not depend on n from the sum:

$$P(r) = \frac{(\epsilon RT)^r e^{-RT}}{r!} \sum_{n=r}^{\infty} \frac{((1 - \epsilon)RT)^{n-r}}{(n-r)!}$$

Now substituting $m = n - r$:

$$P(r) = \frac{(\epsilon RT)^r e^{-RT}}{r!} \sum_{m=0}^{\infty} \frac{((1 - \epsilon)RT)^m}{m!}$$

Recognizing the Taylor series expansion for e^x :

$$e^x = \sum_{m=0}^{\infty} \frac{x^m}{m!}$$

We can rewrite the sum as:

$$P(r) = \frac{(\epsilon RT)^r e^{-RT}}{r!} e^{(1-\epsilon)RT}$$

$$P(r) = \frac{(\epsilon RT)^r e^{-\epsilon RT}}{r!}$$

Which is the same result that we found earlier. The probability of getting r counts from the detector in T seconds is then:

$$\boxed{P(r; \epsilon RT) = \frac{(\epsilon RT)^r e^{-\epsilon RT}}{r!}}$$

(c) We are given $RT = 100$ and $\epsilon = 0.10$. Punching each of these expressions into wolfram alpha I get:

- Probability of having 110 particles in this time:

$$P(n = 110; RT = 100) = \frac{(100)^{110} e^{-100}}{110!} \approx 0.0234$$

- Probability of observing 11 in your detector:

$$P(r = 11; \epsilon RT = 10) = \frac{(10)^{11} e^{-10}}{11!} \approx 0.114$$

- Probability of observing between 90 and 110 particles (inclusive) in this time:

$$P(90 \leq n \leq 110; RT = 100) = \sum_{n=90}^{110} \frac{(100)^n e^{-100}}{n!} \approx 0.7065$$

- Probability of observing between 9 and 11 in your detector:

$$P(9 \leq r \leq 11; \epsilon RT = 10) = \sum_{r=9}^{11} \frac{(10)^r e^{-10}}{r!} \approx 0.364$$

Problem 7

Suppose you have a market of many stocks. The price of a stock at year $i + 1$ is related to the price at the previous year as

$$P_{i+1} = P_i + (P_{i+1} - P_i) = P_i \left(1 + \frac{P_{i+1} - P_i}{P_i}\right) \equiv P_i(1 + r_i)$$

where r_i is the return in year i . If you histogram r_i for many years for a given stock, it will have a distribution with some mean μ and variance σ^2 . Often this is modeled as a normal distribution $\mathcal{N}(\mu, \sigma^2)$. The variance of any given stock is quite high, typically $\sigma \approx 0.3$ (ie 30%). (In finance, this is almost always called the risk, which can be a bit confusing. But in that field, risk = volatility = σ .)

Being an enterprising graduate student, you think "I know how to combine n random variables, take their mean, and drive the variance (=risk) down by increasing n ". Let's model this (diversified) portfolio of stocks as spreading your money equally among n stocks (it's easy to extend to the case of different weights for each stock).

Let stock S_i , $i = 1, \dots, n$ have a return that is a random variable X_i distributed as $\mathcal{N}(\mu_i, \sigma_i^2)$. Form your portfolio by dividing your money into n equal parts invested in these n stocks. Each will contribute a return weighted by $1/n$, and so the sum will have returns which are randomly distributed as:

$$R = \frac{1}{n} \sum_{i=1}^n X_i$$

We want to calculate σ_R = risk of this portfolio. However, you realize that the X_i are not independent. The price of stocks may have a component that affects most in the same way due to variables common to them all, e.g. interest rate changes, global pandemics, oil price shocks, etc. So the covariance matrix V for the n stocks is certainly not diagonal.

Show that you can break up σ_R into two parts, one based upon the variances of the individual stocks and the other based on the covariances. In particular, show that

$$\sigma_R^2 = \frac{1}{n}\bar{\sigma} + \frac{n-1}{n}\bar{c}$$

where $\bar{\sigma}$ = average value of the individual variances,

$$\bar{\sigma} = \frac{1}{n} \sum_i \sigma_i^2$$

and \bar{c} = average value of the covariances

$$\bar{c} = \frac{2}{n(n-1)} \sum_{i < j} V(i, j)$$

This shows that you can diversify away the idiosyncratic risk of individual stocks, as that component decreases with n , but you can't get rid of the systematic/market risk that affects the market as a whole. So, you don't get rewarded with higher expected returns for taking idiosyncratic risk that everyone knows how to minimize, just for the market risk that cannot be diversified away.

Variance of Linear Combination of Random Variables ([Proofwiki](#)) is given by:

$$\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X, Y)$$

So here with our returns defined as:

$$R = \frac{1}{n} \sum_{i=1}^n X_i$$

Then we have:

$$\text{Var}(R) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right)$$

from Bienaymé's Identity ([Wikipedia](#)) we have:

$$\text{Var}(R) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}[X_i] + \frac{2}{n^2} \sum_{i,j=1, i < j}^n \text{Cov}(X_i, X_j)$$

then using the given definitions for average variance and average covariance we have:

$$\bar{\sigma} = \frac{1}{n} \sum_{i=1}^n \sigma_i^2 \Rightarrow \sum_{i=1}^n \text{Var}[X_i] = n\bar{\sigma}$$

and

$$\bar{c} = \frac{2}{n(n-1)} \sum_{i < j} V(i, j) \Rightarrow \sum_{i < j} V(i, j) = \frac{n(n-1)}{2} \bar{c}$$

And since $\text{Var}(X_i) = \sigma_i^2$ and $\text{Cov}(X_i, X_j) = V(i, j)$ we can rewrite the variance of R as:

$$\text{Var}(R) = \frac{1}{n^2}n\bar{v} + \frac{2}{n^2} \cdot \frac{n(n-1)}{2}\bar{c}$$

finally (where $\text{Var}(R) = \sigma_R^2$):

$$\sigma_R^2 = \frac{1}{n}\bar{v} + \frac{n-1}{n}\bar{c}$$

Problem 8

Suppose you have a photon detector that is made of sensor strips (rather than pixels as you might have on a camera), ie they are long in one dimension compared to the other. The detector electronics outputs only a simple binary 1/0 for registering a photon hit or not. Let the strip spacing be l , with strip boundaries at, say, $x = 0, l, 2l, \dots$

You have a source of photons which is uniform over the surface of your detector. Consider hits on a particular strip in location $[l_0, l_0 + l]$.

- (a) What is the probability density function for the x position of the hits in this strip?
- (b) What is $\mathbb{E}(x)$?
- (c) What is $V(x)$, or σ_x ? This is a quite commonly useful result to remember.

- (a) Here the probability distribution is **uniform** over the interval $[l_0, l_0 + l]$. The probability density function (pdf) for a uniform distribution over an interval $[a, b]$ is given by:

$$p(x) = \frac{1}{b-a} \quad \text{for } a \leq x \leq b$$

In our case, $a = l_0$ and $b = l_0 + l$. Therefore, the pdf for the x position of the hits in this strip is:

$$p(x) = \frac{1}{l} \quad \text{for } l_0 \leq x \leq l_0 + l$$

and 0 otherwise.

and you can get $p(x)$ from scratch by normalizing (i.e. forcing area under curve to equal 1) the uniform distribution (constant value between interval) over the interval $[l_0, l_0 + l]$. Area equals base times height, and base equals l , so for area to equal 1, height must equal $\frac{1}{l}$.

- (b) The expected value (mean) $\mathbb{E}(x)$ for a uniform distribution over the interval $[a, b]$ is given by:

$$\mathbb{E}(x) = \frac{a+b}{2}$$

so here we have:

$$\mathbb{E}(x) = \frac{l_0 + (l_0 + l)}{2} = \frac{2l_0 + l}{2} = l_0 + \frac{l}{2}$$

or from scratch can be calculated from $\mathbb{E}(x) = \int_{-\infty}^{\infty} xp(x)dx$:

$$\mathbb{E}(x) = \int_{l_0}^{l_0+l} x \cdot \frac{1}{l} dx$$

$$\mathbb{E}(x) = \frac{1}{l} \cdot \left[\frac{x^2}{2} \right]_{l_0}^{l_0+l} = \frac{1}{l} \cdot \left(\frac{(l_0+l)^2}{2} - \frac{l_0^2}{2} \right)$$

$$\mathbb{E}(x) = \frac{1}{l} \cdot \frac{(l_0^2 + 2l_0l + l^2 - l_0^2)}{2} = \frac{1}{l} \cdot \frac{2l_0l + l^2}{2} = l_0 + \frac{l}{2}$$

(c) The variance $V(x)$ for a uniform distribution over the interval $[a, b]$ is given by:

$$V(x) = \frac{(b-a)^2}{12}$$

so here we have:

$$V(x) = \frac{(l_0 + l - l_0)^2}{12} = \frac{l^2}{12}$$

or from scratch can be calculated from $V(x) = \mathbb{E}(x^2) - (\mathbb{E}(x))^2$:

First we need to calculate $\mathbb{E}(x^2)$:

$$\mathbb{E}(x^2) = \int_{l_0}^{l_0+l} x^2 \cdot \frac{1}{l} dx$$

$$\mathbb{E}(x^2) = \frac{1}{l} \cdot \left[\frac{x^3}{3} \right]_{l_0}^{l_0+l} = \frac{1}{l} \cdot \left(\frac{(l_0+l)^3}{3} - \frac{l_0^3}{3} \right)$$

$$\mathbb{E}(x^2) = \frac{1}{l} \cdot \frac{(l_0^3 + 3l_0^2l + 3l_0l^2 + l^3 - l_0^3)}{3} = \frac{1}{l} \cdot \frac{3l_0^2l + 3l_0l^2 + l^3}{3} = l_0^2 + l_0l + \frac{l^2}{3}$$

Now we can calculate the variance:

$$V(x) = \mathbb{E}(x^2) - (\mathbb{E}(x))^2 = (l_0^2 + l_0l + \frac{l^2}{3}) - (l_0 + \frac{l}{2})^2$$

$$V(x) = (l_0^2 + l_0l + \frac{l^2}{3}) - (l_0^2 + l_0l + \frac{l^2}{4}) = \frac{l^2}{3} - \frac{l^2}{4} = \frac{4l^2 - 3l^2}{12} = \frac{l^2}{12}$$

Problem 9

The χ^2 Distribution: Let $X_i, i = 1, \dots, n$ be n independent normal random variables with Gaussian p.d.f.s $N(\mu_i, \sigma_i^2)$.

Define

$$\chi^2 = \sum_{i=1}^n \left(\frac{x_i - \mu_i}{\sigma_i} \right)^2$$

Then (you can take this as a given) the pdf for χ^2 is

$$p(\chi^2; n) = \frac{1}{2^{\frac{n}{2}} \Gamma(n/2)} (\chi^2)^{\frac{n}{2}-1} e^{-\frac{1}{2}\chi^2}$$

which is called the χ^2 distribution with n degrees of freedom.

- (a) Make plots of this for $n=1, 2, 3$ and 10.
 - (b) Find the mode of this distribution as a function of n .
 - (c) Calculate the mean as a function n .
 - (d) Calculate the variance as a function of n .
-

- (a) Plotted (see figure 1.1) using python with help from gh copilot code completion:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.special import gamma
import scienceplots

plt.style.use(['science', 'notebook'])

n = [1, 2, 3, 10]
x_square = np.linspace(0, 15, 1000)

for i in range(len(n)):
    p_chi_square = (1/(2**(n[i]/2)*gamma(n[i]/2)))*(x_square**(n[i]/2-1))*np.exp(-x_square/2)
    plt.plot(x_square, p_chi_square, label=f'n={n[i]}')

plt.ylim(0, 1)
plt.xlabel(r'$\chi^2$')
plt.ylabel(r'$p(\chi^2;n)$')
plt.legend()
plt.savefig('hw2-p9a.png', dpi=300)
plt.show()
```

- (b) The mode is where the pdf is at it's max. We can get these values from the plot easily but to solve it analytically, we can use the derivative of the pdf and see where it equals 0. Starting with:

$$p(\chi^2; n) = \frac{1}{2^{\frac{n}{2}}\Gamma(n/2)} (\chi^2)^{\frac{n}{2}-1} e^{-\frac{1}{2}\chi^2}$$

$$\frac{dp(\chi^2; n)}{d\chi^2} = \frac{1}{2^{\frac{n}{2}}\Gamma(n/2)} \frac{d \left[(\chi^2)^{\frac{n}{2}-1} e^{-\frac{1}{2}\chi^2} \right]}{d\chi^2}$$

Setting the derivative part to 0 using MATLAB symbolic differentiation:

```
% Define symbolic variables
syms chi_sq n positive

% Define the expression (without normalization constant)
expr = (chi_sq)^(n/2 - 1) * exp(-chi_sq/2);
```

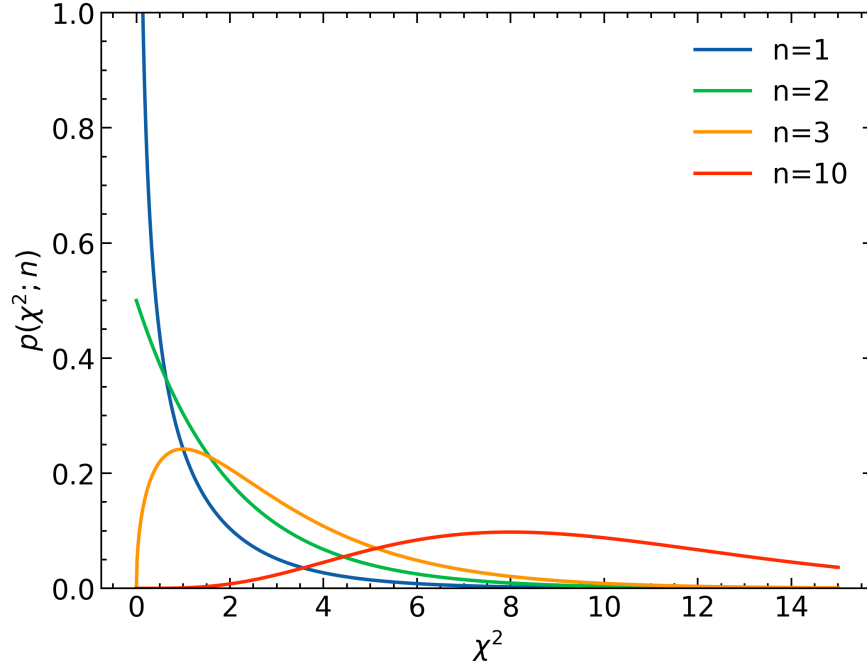


Figure 1.1: Chi-squared distribution for $n=1, 2, 3$, and 10 degrees of freedom.

```
% Take derivative and solve for maximum
derivative = diff(expr, chi_sq);
[mode, params, conds] = solve(derivative == 0, chi_sq, 'ReturnConditions', true);

% Display result
fprintf('%s\n', latex(simplify(mode)));
fprintf('Conditions: %s\n', latex(conds));
```

Result:

$$\chi_{\text{mode}}^2 = \begin{cases} n-2 & \text{for } n \geq 2 \\ 0 & \text{for } n < 2 \end{cases}$$

(c) To calculate the mean as a function of n , we can use the definition of expected value:

$$\mathbb{E}(\chi^2) = \int_0^\infty \chi^2 \cdot p(\chi^2; n) d\chi^2$$

where we have the given pdf

$$p(\chi^2; n) = \frac{1}{2^{\frac{n}{2}} \Gamma(n/2)} (\chi^2)^{\frac{n}{2}-1} e^{-\frac{1}{2}\chi^2}$$

To save time I am going to use MATLAB generated code to solve this step:

```
% Define symbolic variables
```

```

syms chi_sq n positive

% Define chi-squared PDF
pdf = (chi_sq)^(n/2 - 1) * exp(-chi_sq/2) / (2^(n/2) * gamma(n/2));

% Mean
mean_chi = int(chi_sq * pdf, chi_sq, 0, inf);
mean_chi = simplify(mean_chi);

% Display LaTeX output
fprintf('Mean: %s\n', latex(mean_chi));

```

Result:

$$\boxed{\mathbb{E}(\chi^2) = n}$$

(d) To calculate the variance as a function of n , we can use the definition of variance:

$$V(\chi^2) = \mathbb{E}((\chi^2)^2) - (\mathbb{E}(\chi^2))^2$$

where

$$\mathbb{E}(\chi^2) = \int_0^\infty \chi^2 \cdot p(\chi^2; n) d\chi^2$$

and

$$\mathbb{E}((\chi^2)^2) = \int_0^\infty (\chi^2)^2 \cdot p(\chi^2; n) d\chi^2$$

where again we have the given pdf

$$p(\chi^2; n) = \frac{1}{2^{\frac{n}{2}} \Gamma(n/2)} (\chi^2)^{\frac{n}{2}-1} e^{-\frac{1}{2}\chi^2}$$

to save time I am going to again use MATLAB generated code to solve this step:

```

% Define symbolic variables
syms chi_sq n positive

% Define chi-squared PDF
pdf = (chi_sq)^(n/2 - 1) * exp(-chi_sq/2) / (2^(n/2) * gamma(n/2));

% Mean
mean_chi = int(chi_sq * pdf, chi_sq, 0, inf);
mean_chi = simplify(mean_chi);

% Second moment
second_moment = int(chi_sq^2 * pdf, chi_sq, 0, inf);
second_moment = simplify(second_moment);

% Variance

```

```

variance = simplify(second_moment - mean_chi^2);

% Display LaTeX output
fprintf('Mean: %s\n', latex(mean_chi));
fprintf('Second moment: %s\n', latex(second_moment));
fprintf('Variance: %s\n', latex(variance));

```

Result:

$$V(\chi^2) = 2n$$

Problem 10

The Exponential Distribution. The exponential probability density function is

$$E(x; \beta) = \frac{1}{\beta} e^{-\frac{x}{\beta}}, \quad x \geq 0$$

- What is the mode of this distribution?
- Derive the mean.
- Derive the variance.
- Derive the median.
- Calculate the third moment of z and use this and the previous results to derive the standard skewness $\mathbb{E}\left(\left(\frac{x-\mu}{\sigma}\right)^3\right)$.

- For mode we can see that the exponential distribution is a decreasing function for $x \geq 0$, so the maximum value occurs at the left endpoint of the domain, which is $x = 0$. Thus, the mode of the distribution is:

$$\text{Mode} = 0$$

- To derive the mean, we use the definition of the expected value:

$$\mathbb{E}(x) = \int_0^{\infty} x \cdot E(x; \beta) dx = \int_0^{\infty} x \cdot \frac{1}{\beta} e^{-\frac{x}{\beta}} dx$$

```

syms x beta positive

% Define exponential PDF
E_x = (1/beta) * exp(-x/beta);

% Mean
mean_x = int(x * E_x, x, 0, inf);
mean_x = simplify(mean_x);

% Display
fprintf('%s\n', latex(mean_x));

```

$$\mathbb{E}(x) = \beta$$

(c) To derive the variance we use the definition of the variance:

$$V(x) = \mathbb{E}(x^2) - (\mathbb{E}(x))^2$$

where

$$\mathbb{E}(x^2) = \int_0^\infty x^2 \cdot E(x; \beta) dx = \int_0^\infty x^2 \cdot \frac{1}{\beta} e^{-\frac{x}{\beta}} dx$$

and we already have $\mathbb{E}(x) = \beta$ from part (b).

```
syms x beta positive

% Define exponential PDF
E_x = (1/beta) * exp(-x/beta);

% Second moment
E_x2 = int(x^2 * E_x, x, 0, inf);
E_x2 = simplify(E_x2);

% Variance (using mean = beta from part b)
variance = simplify(E_x2 - beta^2);

% Display
fprintf('Second moment: %s\n', latex(E_x2));
fprintf('Variance: %s\n', latex(variance));
```

$$V(x) = \beta^2$$

(d) To derive the median, we need to find the value m such that:

$$\int_0^m E(x; \beta) dx = 0.5$$

```
syms x beta m positive

% Define exponential PDF
E_x = (1/beta) * exp(-x/beta);

% Compute CDF from 0 to m
CDF = int(E_x, x, 0, m);
CDF = simplify(CDF);

% Solve CDF = 0.5 for median
median_val = solve(CDF == 0.5, m);
median_val = simplify(median_val);

% Display
```

```
fprintf('CDF: %s\n', latex(CDF));
fprintf('Median: %s\n', latex(median_val));
```

$$\boxed{\text{Median} = \beta \ln(2)}$$

(e) To calculate the third moment of x , we use the definition of the expected value:

$$\mathbb{E}(x^3) = \int_0^\infty x^3 \cdot E(x; \beta) dx = \int_0^\infty x^3 \cdot \frac{1}{\beta} e^{-\frac{x}{\beta}} dx$$

Then, to derive the standard skewness, we use the formula:

$$\mathbb{E}((x - \mu)^3) = \int_0^\infty (x - \mu)^3 \cdot E(x; \beta) dx = \int_0^\infty (x - \mu)^3 \cdot \frac{1}{\beta} e^{-\frac{x}{\beta}} dx$$

```
syms x beta positive

% Define exponential PDF
E_x = (1/beta) * exp(-x/beta);

% Third raw moment
E_x3 = int(x^3 * E_x, x, 0, inf);
E_x3 = simplify(E_x3);

% Third central moment (using mu = beta from part b)
mu = beta;
third_central = int((x - mu)^3 * E_x, x, 0, inf);
third_central = simplify(third_central);

% Standard skewness (using sigma = beta from part c)
sigma = beta;
skewness = simplify(third_central / sigma^3);

% Display
fprintf('Third moment: %s\n', latex(E_x3));
fprintf('Third central moment: %s\n', latex(third_central));
fprintf('Skewness: %s\n', latex(skewness));
```

$$\boxed{\text{Skewness} = 2}$$