MATH 368/621 Fall 2020 Homework #2

Frank Palma Gomez

Tuesday 22nd September, 2020

Problem 1

These exercises will introduce review expectation and variance and introduce covariance as well as expectation and variance of multidimensional (vector) r.v.'s.

(a) [harder] Consider a sequence of independent r.v.'s $X_1, \ldots, X_n$ and prove that

$$\mathbb{E}\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} \mathbb{E}\left[X_i\right].$$

$$\mathbb{E}\left[\prod_{i=1}^{n} X_i\right] = \mathbb{E}\left[X_1 \cdot, \ldots, \cdot, X_n\right] = \mathbb{E}\left[X_1\right] \cdot \ldots \cdot \mathbb{E}\left[X_n\right] = \prod_{i=1}^{n} \mathbb{E}\left[X_i\right]$$

(b) [easy] Prove that $\mathbb{C}\text{ov}\left[X_1, X_2\right] = \mathbb{E}\left[(X_1 - \mu_1)(X_2 - \mu_2)\right]$.

$$\begin{aligned}
\mathbb{C}\text{ov}\left[X_1, X_2\right] &= \mathbb{E}\left[(X_1 - \mu_1)(X_2 - \mu_2)\right] \\
&= \mathbb{E}\left[X_1 X_2 - \mu_1 X_1 - \mu_2 X_2 + \mu_1 \mu_2\right] \\
&= \mathbb{E}\left[X_1 X_2\right] - \mu_1 \mu_2
\end{aligned}$$

(c) [easy] Prove that $\mathbb{C}\text{ov}\left[X, X\right] = \mathbb{V}\text{ar}\left[X\right]$.

$$\begin{aligned}
\mathbb{C}\text{ov}\left[X, X\right] &= \mathbb{E}\left[XX\right] - \mu\mu \\
&= \mathbb{E}\left[X\right] \mathbb{E}\left[X\right] - \mu\mu \\
&= \mathbb{E}\left[X - \mu\right] \mathbb{E}\left[X - \mu\right] \\
&= \mathbb{E}\left[(X - \mu^2\right] \\
&= \mathbb{V}\text{ar}\left[X\right]
\end{aligned}$$

(d) [easy] Prove that $\mathbb{C}\mathrm{ov}\left[X_1, X_2\right] = \mathbb{C}\mathrm{ov}\left[X_2, X_1\right]$.

$$\begin{aligned}
\mathbb{C}\mathrm{ov}\left[X_1, X_2\right] &= \mathbb{E}\left[X_1\right]\mathbb{E}\left[X_2\right]\mu_1\mu_2 \\
&= \mathbb{E}\left[X_2\right]\mathbb{E}\left[X_1\right] - \mu_2\mu_1 \\
&= \mathbb{C}\mathrm{ov}\left[X_2, X_1\right]
\end{aligned}$$

(e) [easy] Prove that $\mathbb{C}\mathrm{ov}\left[a_1X_1, a_2X_2\right] = a_1a_2\mathbb{C}\mathrm{ov}\left[X_1, X_2\right]$.

$$\begin{aligned}
\mathbb{C}\mathrm{ov}\left[a_1X_1, a_2X_2\right] &= \mathbb{E}\left[(a_1X_1 - a_1\mu_1)(a_2X_2 - a_2\mu_2)\right] \\
&= \mathbb{E}\left[a_1a_2(X_1 - \mu_1)(X_2 - \mu_2)\right] \\
&= a_1a_2\mathbb{E}\left[(X_1 - \mu_1)(X_2 - \mu_2)\right] \\
&= a_1a_2\mathbb{C}\mathrm{ov}\left[X_1, X_2\right]
\end{aligned}$$

(f) [easy] Prove that $\mathbb{C}\mathrm{ov}\left[X_1 + X_3, X_2\right] = \mathbb{C}\mathrm{ov}\left[X_1, X_2\right] + \mathbb{C}\mathrm{ov}\left[X_2, X_3\right]$.

$$\begin{aligned}
\mathbb{C}\mathrm{ov}\left[X_1 + X_3, X_2\right] &= \mathbb{E}\left[(X_1 + X_3 - \mu_1 - \mu_3)(X_2 - \mu_2)\right] \\
&= \mathbb{E}\left[X_1X_2 - \mu_1X_1 + X_3X_2 - \mu_2X_3\right] \\
&= \mathbb{E}\left[X_1X_2\right] - \mu_2\mu_1 + \mathbb{E}\left[X_3, X_2\right] - \mu_2\mu_3 \\
&= \mathbb{C}\mathrm{ov}\left[X_1, X_2\right] + \mathbb{C}\mathrm{ov}\left[X_2, X_3\right]
\end{aligned}$$

(g) [harder] [MA] Prove that

$$\mathbb{C}\mathrm{ov}\left[\sum_{i\in A} X_i, \sum_{j\in B} Y_j\right] = \sum_{i\in A}\sum_{j\in B}\mathbb{C}\mathrm{ov}\left[X_i, Y_j\right]$$

.

$$\begin{aligned}
\mathbb{C}\mathrm{ov}\left[\sum_{i\in A} X_i, \sum_{j\in B} Y_j\right] &= \mathbb{E}\left[(\sum_{i\in A} X_i - \sum_{i\in A}\mu_i)(\sum_{j\in B} Y_j - \sum_{j\in B}\mu_j)\right] \\
&= \mathbb{E}\left[\sum_{i\in A} X_i \sum_{j\in B} Y_i\right] - \sum_{j\in B}\mu_j\sum_{i\in A}\mu_i \\
&= \sum_{i\in A}\sum_{j\in B}\mathbb{E}\left[X_i, Y_j\right] - \mu_i\mu_j \\
&= \sum_{i\in A}\sum_{j\in B}\mathbb{C}\mathrm{ov}\left[X_i, Y_j\right]
\end{aligned}$$

(h) [difficult] Prove that

$$\mathbb{V}\text{ar}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{C}\text{ov}\left[X_i, X_j\right]$$

without using the vector formulas.

$$\begin{aligned}
\mathbb{V}\text{ar}\left[\sum_{i=1}^{n} X_i\right] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \left[\mathbb{E}\left[\sum_{i=1}^{n} X_i\right]\right]^2 \\
&= \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right] - \mathbb{E}\left[\sum_{i=1}^{n} X_i\right]\mathbb{E}\left[\sum_{j=1}^{n} X_j\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}\left[X_i X_j\right] - \mu_i \mu_j \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{C}\text{ov}\left[X_i, X_j\right]
\end{aligned}$$

(i) [easy] Prove $\mathbb{E}\left[a\boldsymbol{X} + \boldsymbol{c}\right] = a\boldsymbol{\mu} + \boldsymbol{c}$ where the following are constants: $a \in \mathbb{R}, \boldsymbol{c} \in \mathbb{R}^K$.

$$\begin{aligned}
\mathbb{E}\left[a\boldsymbol{X} + \boldsymbol{c}\right] &= [a\mathbb{E}\left[X_1\right] + c_1, \ldots, a\mathbb{E}\left[X_n\right] + c_n]^T \\
&= a[\mathbb{E}\left[X_1\right], \ldots, \mathbb{E}\left[X_n\right]]^T + [c_1, \ldots, c_n]^T \\
&= a\boldsymbol{\mu} + \boldsymbol{c}
\end{aligned}$$

(j) [easy] Prove $\mathbb{V}\text{ar}\left[\boldsymbol{c}^\top \boldsymbol{X}\right] = \boldsymbol{c}^\top \Sigma \boldsymbol{c}$ where $\boldsymbol{c} \in \mathbb{R}^K$, a constant and $\Sigma := \mathbb{V}\text{ar}\left[\boldsymbol{X}\right]$, the variance-covariance matrix of the vector r.v. $\boldsymbol{X}$. This is marked easy since it's in the notes.

$$\begin{aligned}
\mathbb{V}\text{ar}\left[\boldsymbol{c}^\top \boldsymbol{X}\right] &= \mathbb{V}\text{ar}\left[c_1 X_1 + \ldots + c_k X_k\right] \\
&= \mathbb{V}\text{ar}\left[X_1 + \ldots + X_k\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{C}\text{ov}\left[c_i X_i, c_j X_j\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \sigma_{ij} \\
&= \boldsymbol{c}^\top \Sigma \boldsymbol{c}
\end{aligned}$$

(k) [easy] Why is $\boldsymbol{c}^\top \Sigma \boldsymbol{c}$ called a "quadratic form?" Read about it on wikipedia.

Its called quadratic form because they are qudratic polonomials in n variables.

Problem 2

These exercises are about the Multinomial distribution.

(a) [easy] Explain in English why $\boldsymbol{B} \sim \text{Multinomial}\,(1,\, \boldsymbol{p})$ is the multidimensional generalization of the Bernoulli r.v.

Binomial $(n,\, p_i)$ is the number of successes after n trials with probability p_i
Since n=1 then binomial converts to a bernoulli. For that reason we could express
Multinomial $(1,\, \boldsymbol{p})$ as a row vector of bernoulli's

(b) [easy] Explain in English why the following should be true. Remember how the sampling from the bag works.

$$\binom{n}{x_1, x_2, \ldots, x_K} = \binom{n}{x_1}\binom{n-x_1}{x_2}\binom{n-(x_1+x_2)}{x_3} \cdot \ldots \cdot \binom{n-(x_1+x_2+\ldots x_{K-1})}{x_K}$$

Say you have n fruits in a bag. The fruits can be denoted as x_1, x_2
If we sample from the bag and get x_1 type fruit,
then the bag has a remainder of $n - x_1$ when we sample again
Additionally, this signifies that sampling is dependent on previous samples

(c) [harder] Prove the combinatorial identity in (b).

$$\binom{n}{x_1, x_2, \ldots, x_K} = \frac{n!}{x_1!(n-x_1)!} \cdot \ldots \cdot \frac{(n - x_1 - \ldots - x_{K-1})!}{x_K!(n - x_1 - \ldots - x_{K-1} - x_K)!}$$

Canceling terms yields to

$$\begin{aligned} \frac{n!}{x_1! \ldots x_K!} &= \frac{n!}{x_1! \ldots x_K!(n - x_1 - \ldots - x_K)!} \\ &= \frac{n!}{x_1! \ldots x_K!} \end{aligned}$$

(d) [easy] Consider the following bag of marbles.

Draw from replacement 37 times. What is the probability of getting 10 red, 17 green, 6 blue and 4 yellow? Compute explicitly to the nearest two significant digits.

$$\binom{37}{10, 17, 6, 4} (\frac{10}{37})^{10} (\frac{17}{37})^{17} (\frac{6}{37})^{6} (\frac{4}{37})^{4} = 0.01$$

(e) [difficult] [MA] If $\boldsymbol{X} \sim \text{Multinomial}\,(n,\, \boldsymbol{p})$, prove that its JMF sums to one, i.e. $\sum_{\boldsymbol{x} \in \text{Supp}[\boldsymbol{X}]} p_{\boldsymbol{X}}(\boldsymbol{x}) = 1$.

(f) [difficult] [MA] If $\boldsymbol{X} \sim \text{Multinomial}\,(n,\, \boldsymbol{p})$, prove that any marginal distribution is binomial with n and p_j as parameters i.e.

$$p_{X_j}(x_j) = \text{Binomial}\,(n,\, p_j)$$

We only assumed this in class because it makes sense conceptual given balls being sampled from an urn, but it was never explicitly proven.

Assume j = 1

$$\begin{aligned}
\mathbb{P}\,(X_1 = x_1) &= \sum_{x_2, \ldots, x_n} (\frac{n!}{x_1!, \ldots, x_k!}) p_1^{x_1} \cdot \ldots \cdot p_K^{x_k} \\
&= \frac{n!}{x_1 (n - x_1)!} p_1^{x_1} \sum_{x_2, \ldots, x_k} \binom{n - x_1}{x_2, \ldots x_k} p_2^{x_2} \cdot \ldots \cdot p_K^{x_k} \\
&= \frac{n!}{x_1 (n - x_1)!} p_1^{x_1} (1 - p_1)^{n - x_1} \\
&= \text{Binomial}\,(n,\, p_1) \\
&= \text{Binomial}\,(n,\, p_j)
\end{aligned}$$

(g) [E.C.] [MA] If $\boldsymbol{X} \sim \text{Multinomial}(n, \boldsymbol{p})$, find the JMF of any subset of $X_1, \ldots, X_k$. Is it technically multinomial? This is not much harder than the previous problem if formulated carefully.

Let $n\prime = n - x_1 - \ldots - x_i$ and $\boldsymbol{p\prime} = \frac{p_{i+1}}{1-p_1-\ldots-p_i} + \ldots + \frac{p_k}{1-p_1-\ldots-p_i}$

$$
\begin{aligned}
\mathbb{P}(X_1 = x_1, \ldots X_i = x_i) &= \sum_{x_1 \ldots x_k} \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k} \\
&= p_1^{x_1} \ldots p_i^{x_i} \binom{n}{x_1! \ldots x_i!} \sum_{x_{i+1} \ldots x_k} \binom{n\prime}{x_{i+1}! \ldots x_k!} p\prime_1^{x_i+1} \ldots p\prime_k^{x_k} \\
&= p_1^{x_1} \ldots p_i^{x_i} \binom{n - \sum_{j=i+1}^{k} x_j}{x_1! \ldots x_i!} \\
&= \text{Multinomial}\left(n - \sum_{j=i+1}^{k} x_j, \boldsymbol{p}\right)
\end{aligned}
$$

Where $\boldsymbol{p} = [\frac{p_1}{1-p_{i+1}-\ldots-p_k}, \ldots, \frac{p_i}{1-p_{i+1}-\ldots-p_k}]$

(h) [harder] Explain in English why the following should be true. Remember how the sampling from the bag works.

$$\boldsymbol{B}_1, \ldots, \boldsymbol{B}_n \overset{iid}{\sim} \text{Multinomial}(1, \boldsymbol{p}) \quad \text{then} \quad \boldsymbol{X} := \sum_{i=1}^{n} \boldsymbol{B}_i \sim \text{Multinomial}(n, \boldsymbol{p})$$

If you have n=1 then you only have 1 trial or you're only sampling once
Adding all these together is like chaining n different trials.
We can do this because they are iid.

(i) [harder] Find the answer by reasoning in English. No need to prove mathematically.

$$\boldsymbol{X}_1, \ldots, \boldsymbol{X}_r \overset{iid}{\sim} \text{Multinomial}(n, \boldsymbol{p}) \quad \text{then} \quad \boldsymbol{T} := \sum_{i=1}^{r} \boldsymbol{X}_i \sim \ ?$$

We know n represent the number of independent trials. If we add r different independent trials then we now have $r \cdot n$ trials. So we have

$$T := \sum_{i=1}^{r} X_i = \text{Multinomial}\left(\sum_{i=1}^{r} n,\, \boldsymbol{p}\right)$$

(j) [easy] If $\boldsymbol{X} \sim \text{Multinomial}(n,\, \boldsymbol{p})$, find $p_{\boldsymbol{X}_{-j}|X_j}(\boldsymbol{x}_{-j}, x_j)$. This is marked easy since it's in the notes.

Let $n' = n - x_j$ and $p' = \frac{p_1}{1-p_j} \dots \frac{p_k}{1-p_j}$

$$\begin{aligned}
\mathbb{P}(X_{i-j}, X_j) &= \frac{\text{Multinomial}(n,\, p)}{\text{Binomial}(n,\, p_j)} \\
&= \frac{\binom{n}{x_1,\dots,x_j,\dots,x_K} p_1^{x_1} \cdot \dots \cdot p_j^{x_j} \cdot \dots p_k^{x_k}}{\binom{n}{x_j} p_j^{x_j} (1-p_j)^{n-x_j}} \\
&= \frac{\frac{n!}{x_1!\dots x_j! \dots x_k!} \mathbb{1}_{\sum_{i=1}^k x_i = n}\, p_1^{x_1} \cdot \dots \cdot p_{j-1}^{x_{j-1}} \cdot \dots p_k^{x_k} \prod_{i=1}^{k} \mathbb{1}_{x_i \in J_n}}{\binom{n}{x_j} (1-p_j)^{n-x_j}} \\
&= \text{Multi}_{k-1}(n', \boldsymbol{p}')
\end{aligned}$$

(k) [E.C.] [MA] If $\boldsymbol{X} \sim \text{Multinomial}(n,\, \boldsymbol{p})$, find a proof for $\mathbb{C}\text{ov}[X_i, X_j] = -np_ip_j$ that is qualitatively different than the one we did in class.

$$\begin{aligned}
\mathbb{C}\text{ov}[X_i, X_j] &= \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\,\mathbb{E}[X_j] \\
&= \mathbb{E}[X_i, X_j] - n^2 p_i p_j
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[X_i X_j] &= \sum x_i x_j \mathbb{P}(X_1 = x_1 \dots X_k = x_k) \\
&= \sum x_i x_j \left(\frac{n!}{x_1! \dots x_k!}\right) p_1^{x_1} \dots p_i^{x_i} \dots p_k^{x_k} \\
&= \sum x_i x_j \frac{n!}{x_1! \dots x_{i-1}! x_{j-1}! \dots x_k!} p_1^{x_1} \cdot \dots \cdot p_i^{x_i - 1} \cdot p_j^{x_j - 1} \dots p_k^{x_k} \\
&= p_i p_j \sum \frac{n(n-1)(n-2)\dots}{x_1! \dots x_{i'}! x_{j'}! \dots x_k!} \\
&= n(n-1) p_i p_j
\end{aligned}$$

$$\begin{aligned}
\mathbb{C}\text{ov}[X_i, X_j] &= \mathbb{E}[X_i, X_j] - n^2 p_i p_j \\
&= n(n-1) p_i p_j - n^2 p_i p_j \\
&= -n p_i p_j
\end{aligned}$$

(l) [harder] If $\boldsymbol{X} \sim \text{Multinomial}(n, \boldsymbol{p})$ where $\dim[\boldsymbol{X}] = K$ and $\boldsymbol{p} = \frac{1}{K}\mathbf{1}_K$. What is the limit of $\mathbb{C}\text{ov}[X_i, X_j]$ as K gets large but n is fixed. Why does this make sense?

$$\lim_{K\to\infty} -np_ip_j = \lim_{K\to\infty} \frac{1}{K^2} = 0$$

This makes sense because as dimensions increase, the correlation between random variables becomes smaller since everything has equal probability.

(m) [easy] Correlation ρ is a unitless measure bounded between $[-1, 1]$ and is a type of normalized covariance metric. It is defined for two r.v.'s as

$$\rho_{1,2} := \mathbb{C}\text{orr}[X_1,\ X_2] := \frac{\sigma_{1,2}}{\sigma_1\sigma_2} = \frac{\mathbb{C}\text{ov}[X_1, X_2]}{\mathbb{SD}[X_1]\,\mathbb{SD}[X_2]} = \frac{\mathbb{C}\text{ov}[X_1, X_2]}{\sqrt{\mathbb{V}\text{ar}[X_1]\,\mathbb{V}\text{ar}[X_2]}}$$

where $\mathbb{SD}[\cdot]$ denotes the standard deviation of a r.v., the square root of its variance. Find $\mathbb{C}\text{orr}[X_i,\ X_j]$ for two arbitrary elements in the r.v. vector $\boldsymbol{X} \sim \text{Multinomial}(n,\ \boldsymbol{p})$.

$$\begin{aligned}
\mathbb{C}\text{orr}[X_1,\ X_2] &= \frac{\mathbb{C}\text{ov}[X_1, X_2]}{\sqrt{\mathbb{V}\text{ar}[X_1]\,\mathbb{V}\text{ar}[X_2]}} \\
&= \frac{\mathbb{E}[X_iX_j] - p_ip_j}{\sqrt{np_i(1-p_i)np_j(1-p_j)}} \\
&= \frac{-p_ip_j}{\sqrt{p_i(1-p_i)p_j(1-p_j)}}
\end{aligned}$$

(n) [easy] If $\boldsymbol{c} = [1\ 2\ 3\ 4]^\top$, compute the inner product $\boldsymbol{c}^\top\boldsymbol{c}$ and the outer product $\boldsymbol{c}\boldsymbol{c}^\top$.

$$\begin{aligned}
\boldsymbol{c}^\top\boldsymbol{c} &= 30 \\
\boldsymbol{c}\boldsymbol{c}^\top &= 20
\end{aligned}$$